A Personalized RNN Model
for
Predicting Weight Change

Jamie Williams
@jamieinfinity

Saint Louis Deep Learning Meetup
Tuesday, January 17, 2017

Local DL Slack Channel

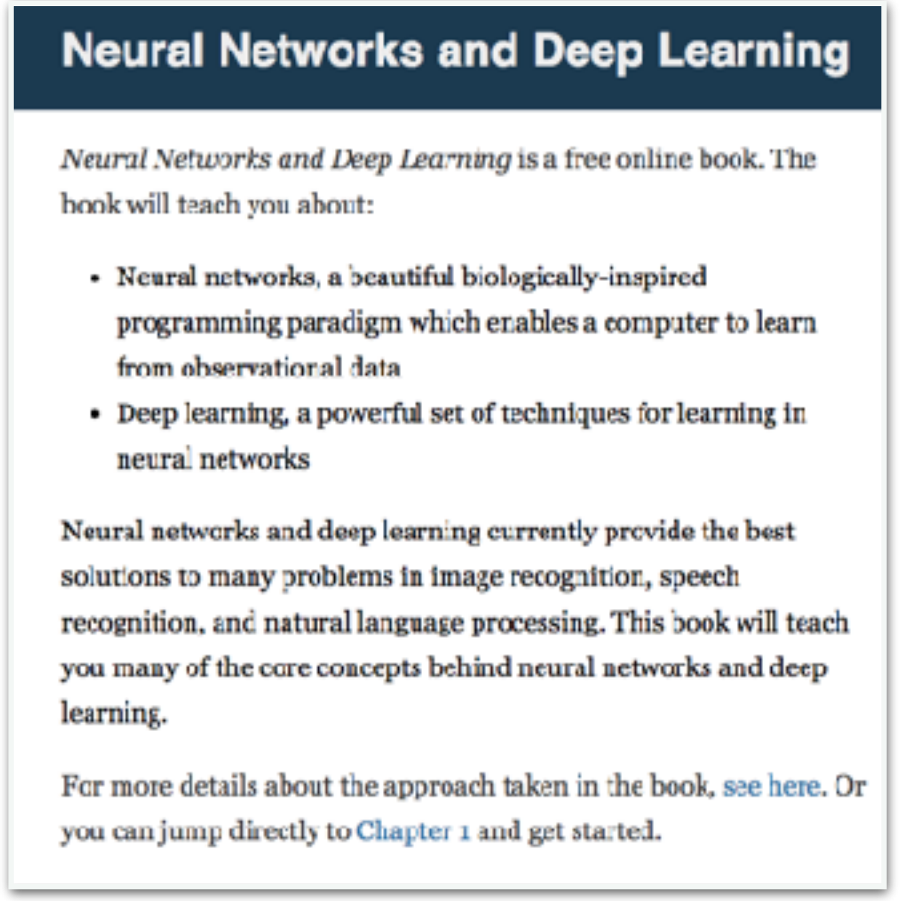
DL channel: #deep_learning

Slack team: https://stl-tech.slack.com/

Signup at: https://stltech.herokuapp.com/

Deep Learning Resources

Michael Nielsen’s site
neuralnetworksanddeeplearning.com



Neural Networks and Deep Learning

Neural Networks and Deep Learning is a free online book. The book will teach you about:

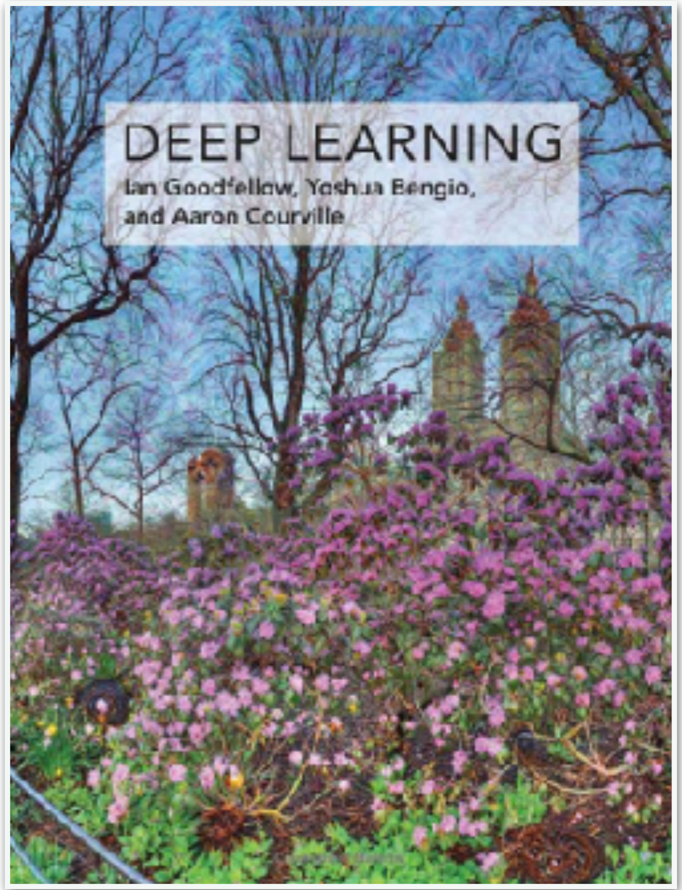
- Neural networks, a beautiful biologically-inspired programming paradigm which enables a computer to learn from observational data
- Deep learning, a powerful set of techniques for learning in neural networks

Neural networks and deep learning currently provide the best solutions to many problems in image recognition, speech recognition, and natural language processing. This book will teach you many of the core concepts behind neural networks and deep learning.

For more details about the approach taken in the book, see here. Or you can jump directly to Chapter 1 and get started.

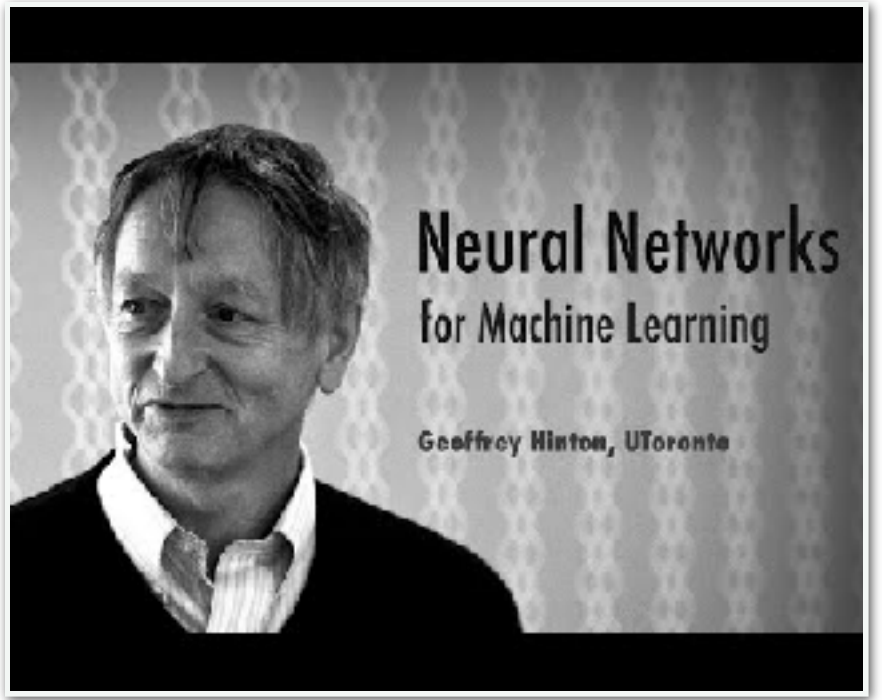
Machine Learning Mastery
machinelearningmastery.com

Deep Learning book
by Goodfellow, Bengio, Courville

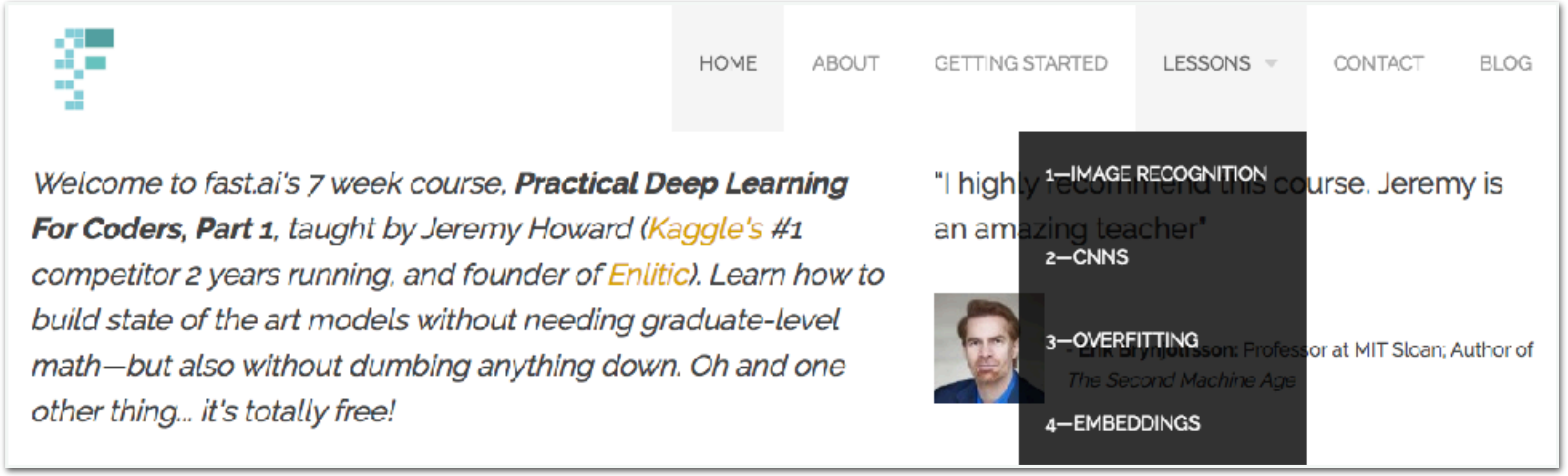


NN for Machine Learning
by G. Hinton

coursera

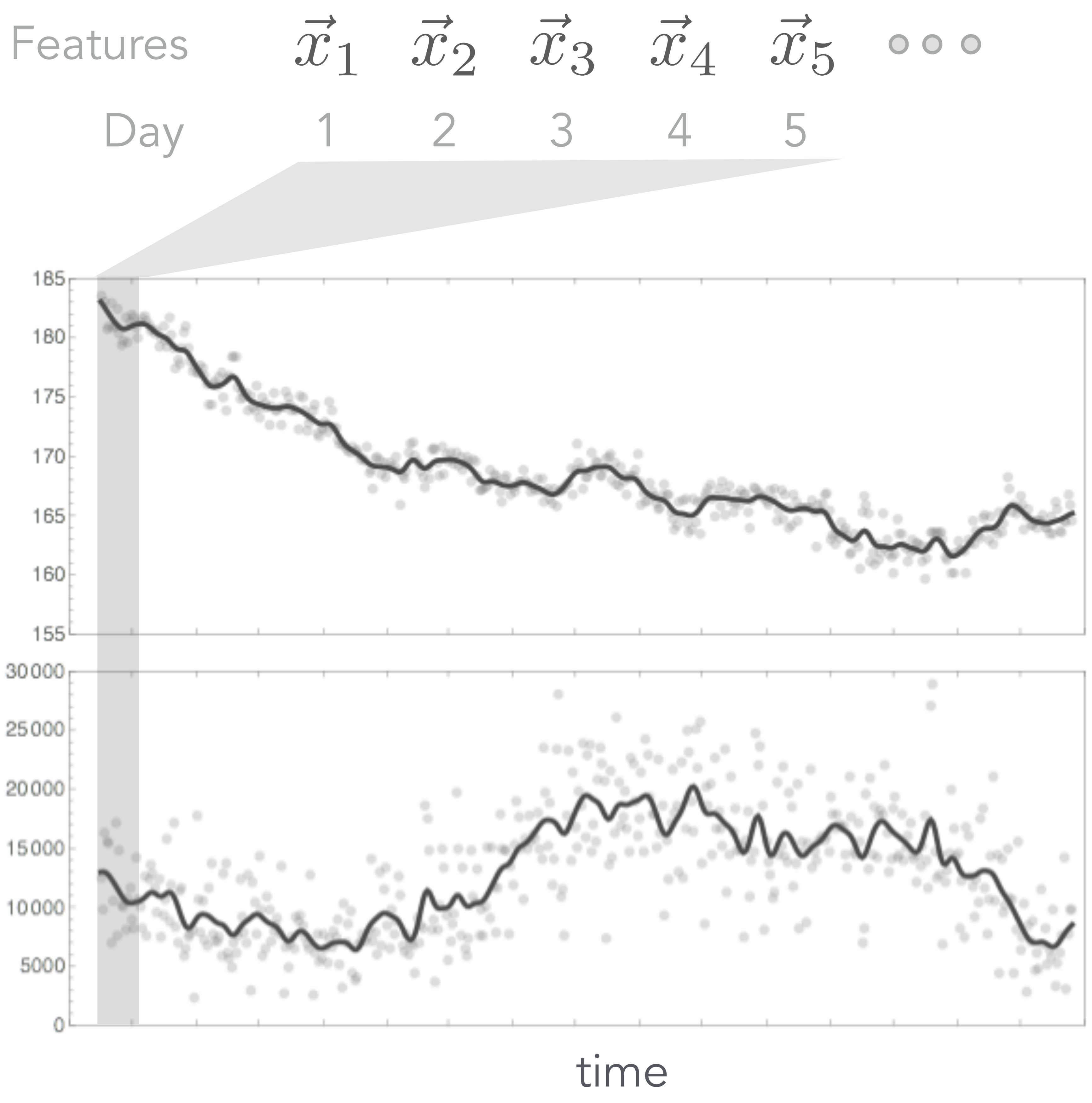


course.fast.ai/index.html

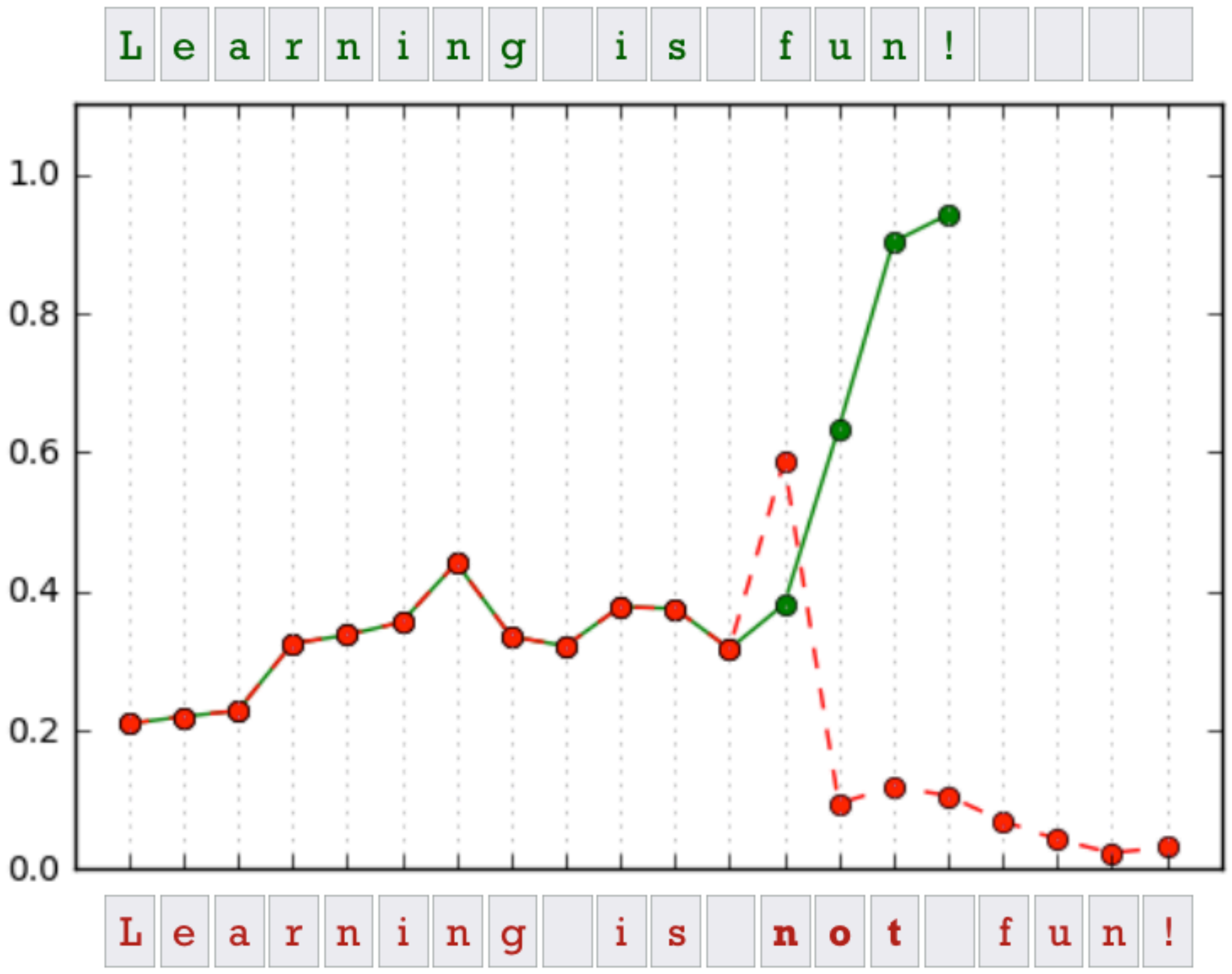


Sequential Data

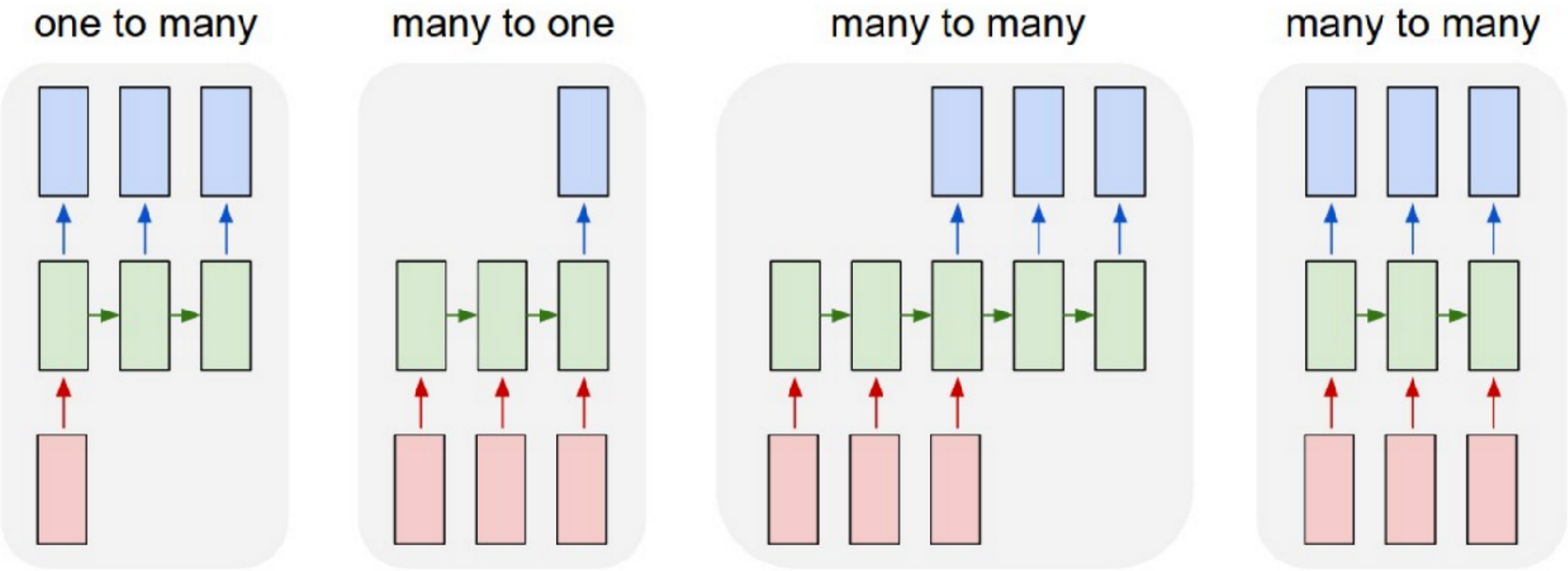
Time Series Prediction



Natural Language Processing



Recurrent Neural Network (RNN) Models

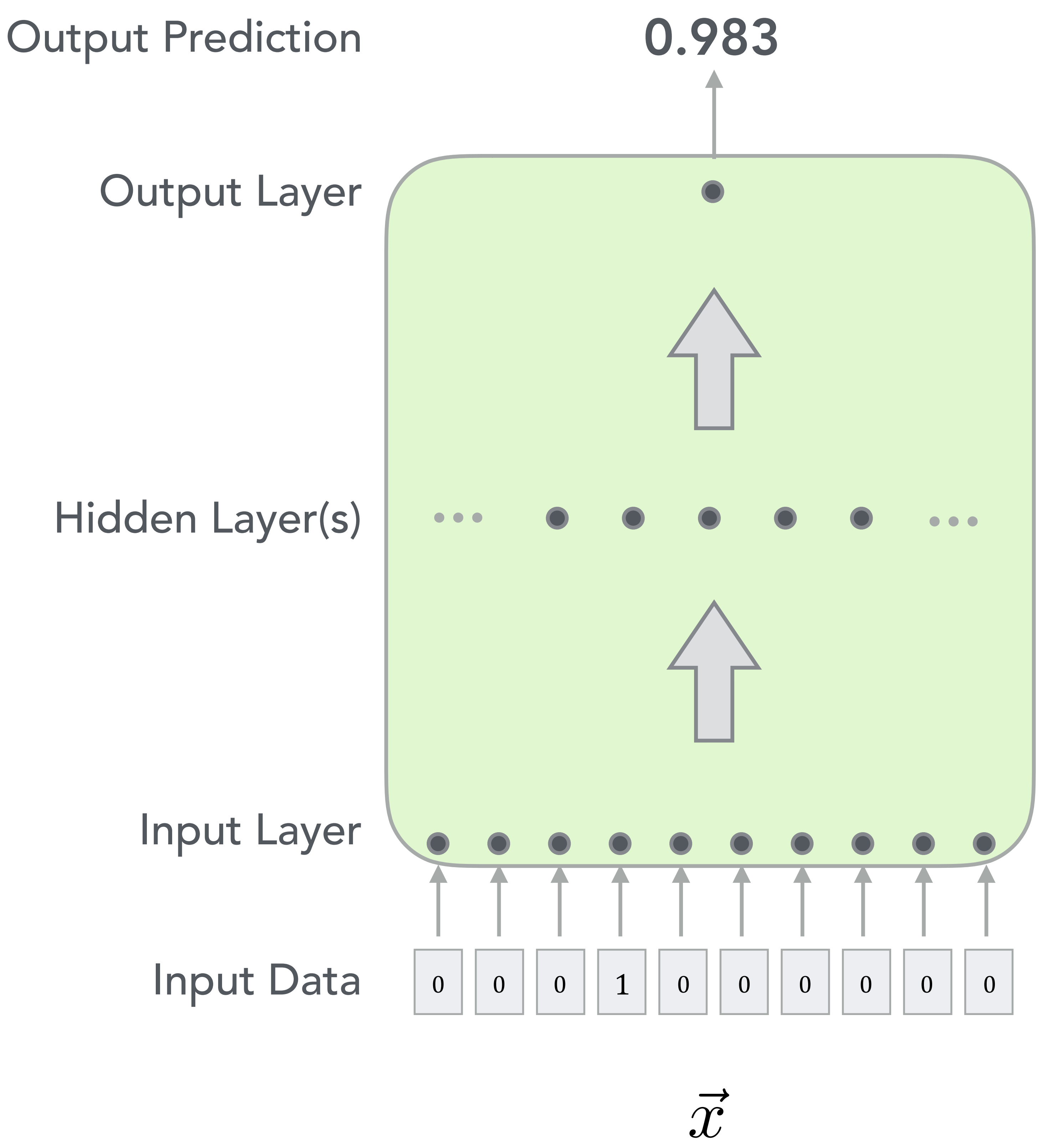


Examples:	Image Captioning	Sentiment Classification	Machine Translation	Video Classification

From: http://karpathy.github.io/2015/05/21/rnn-effectiveness/

Network Architecture

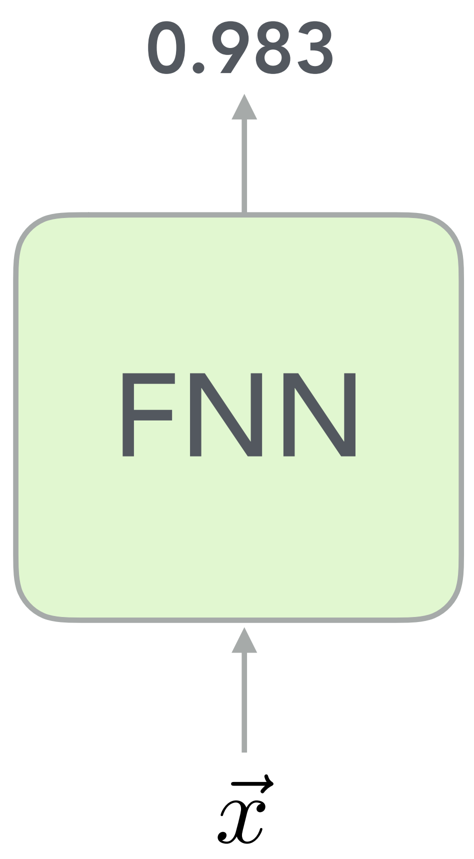
Standard Feedforward Network



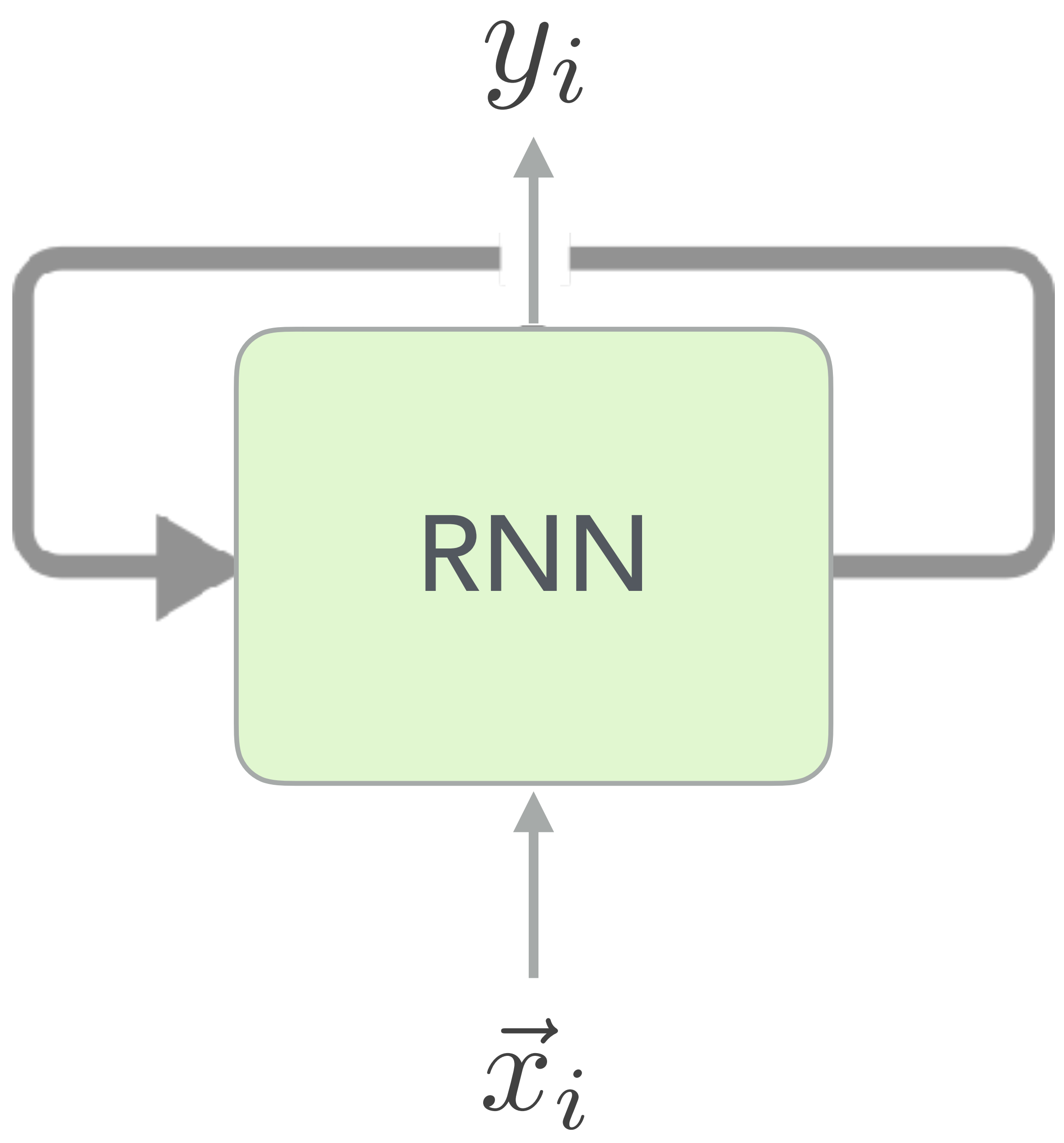
Recurrent Neural Network

Network Architecture

Standard Feedforward Network

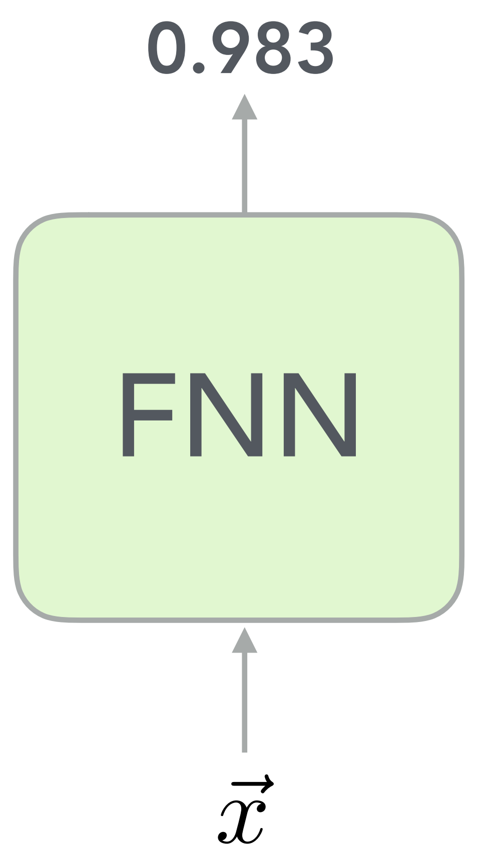


Recurrent Neural Network



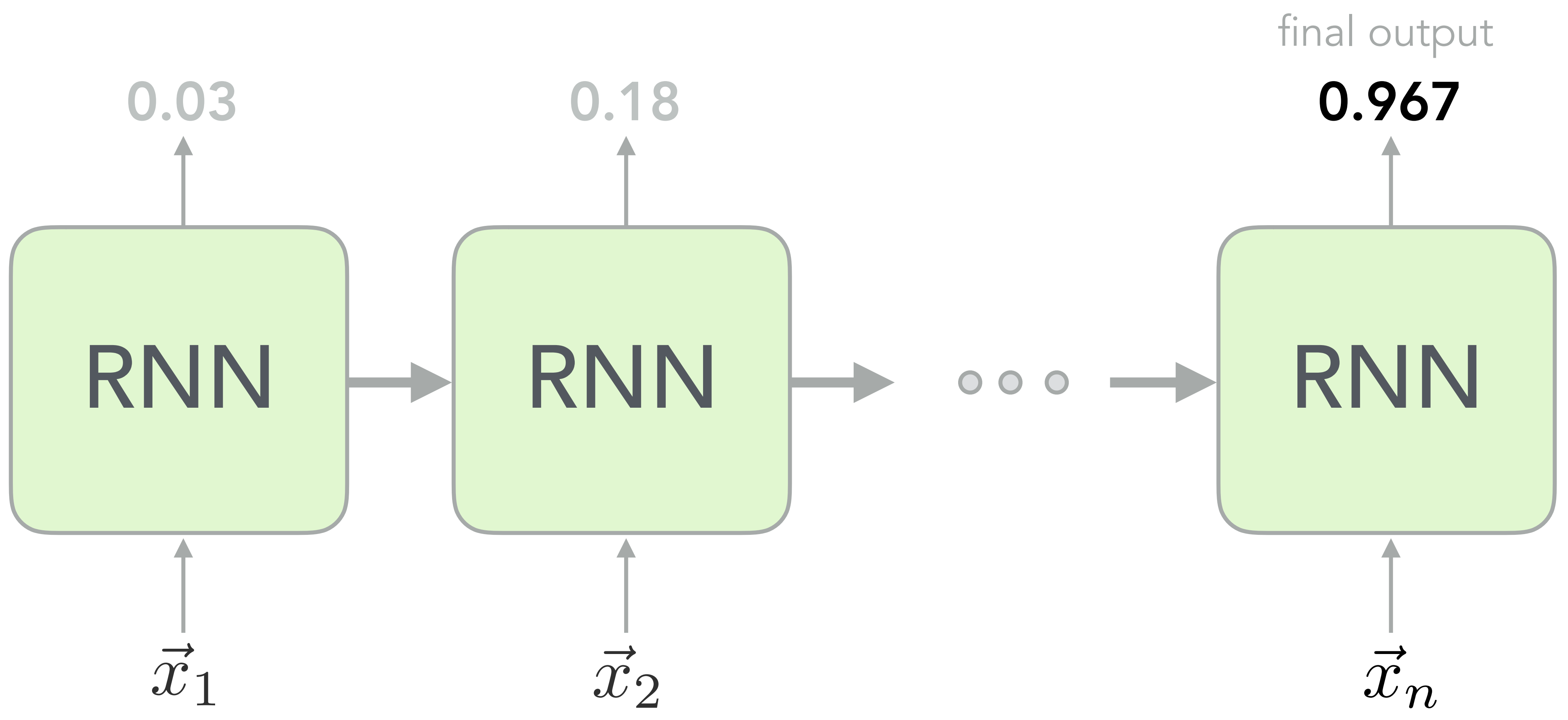
Network Architecture

Standard Feedforward Network



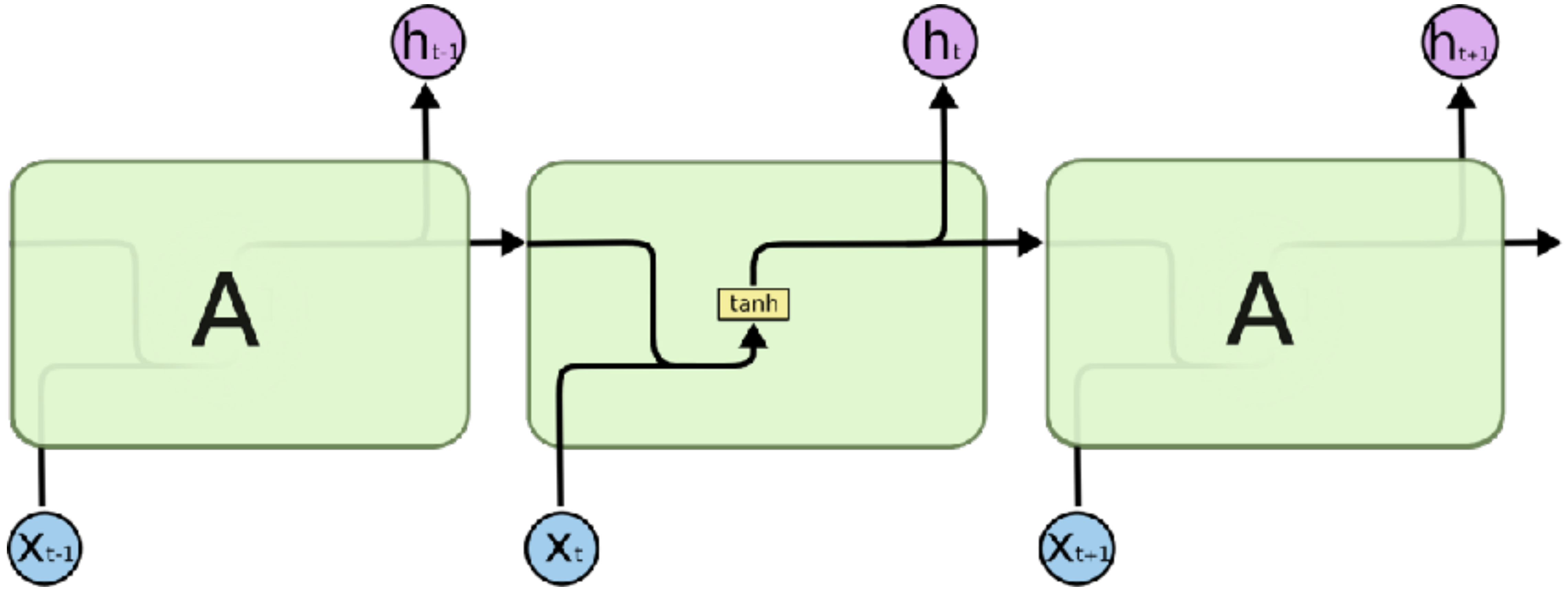
Recurrent Neural Network

(unrolled)



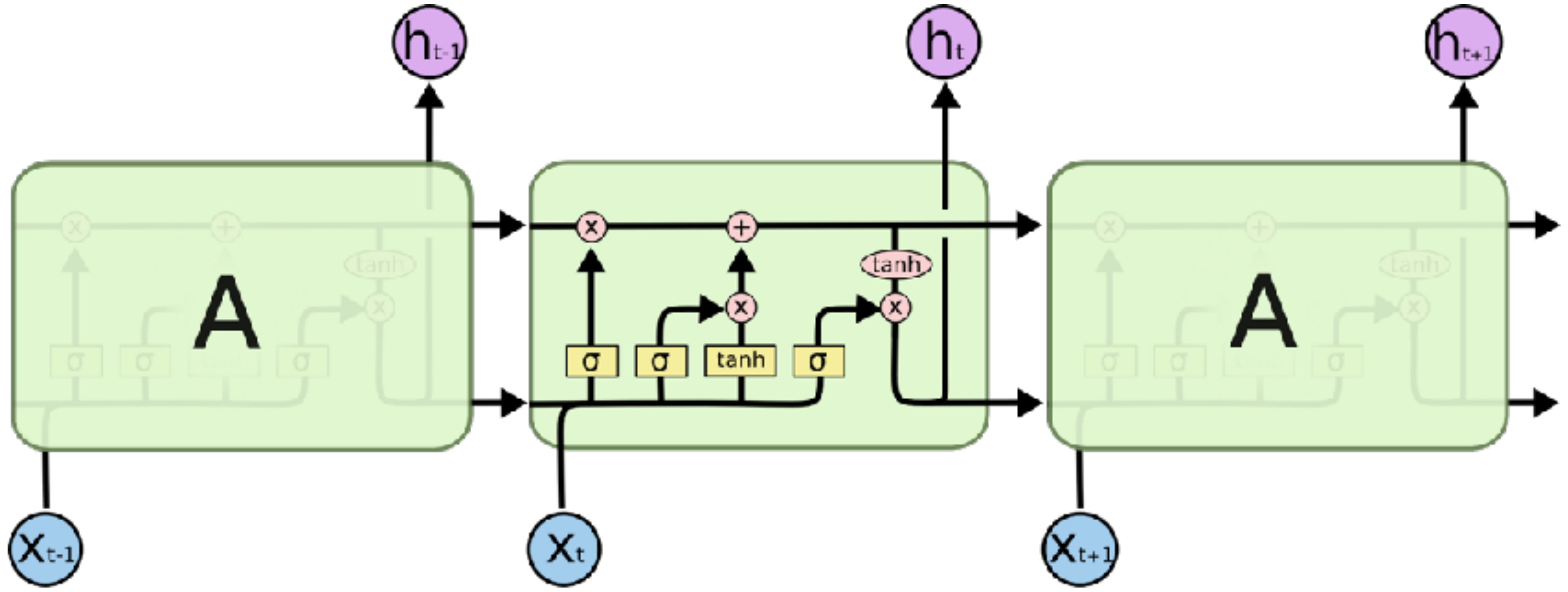
Memory Units

Standard RNN



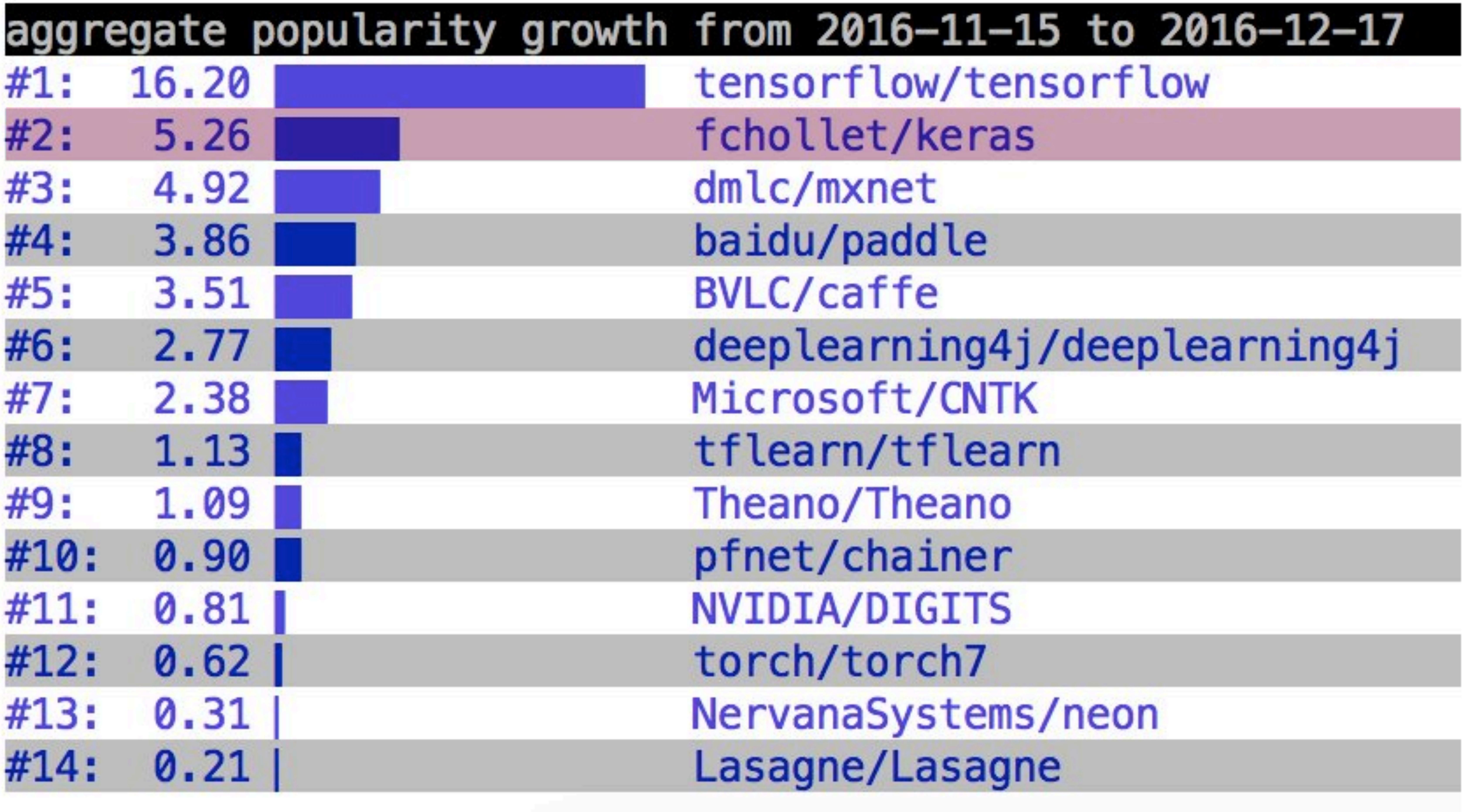
LSTM Units

Long Short-Term Memory



From: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

Deep Learning Tools



aggregate popularity growth from 2016-11-15 to 2016-12-17		
#1:	16.20	tensorflow/tensorflow
#2:	5.26	fchollet/keras
#3:	4.92	dmlc/mxnet
#4:	3.86	baidu/paddle
#5:	3.51	BVLC/caffe
#6:	2.77	deeplearning4j/deeplearning4j
#7:	2.38	Microsoft/CNTK
#8:	1.13	tflearn/tflearn
#9:	1.09	Theano/Theano
#10:	0.90	pfnet/chainer
#11:	0.81	NVIDIA/DIGITS
#12:	0.62	torch/torch7
#13:	0.31	NervanaSystems/neon
#14:	0.21	Lasagne/Lasagne

keras

https://twitter.com/fchollet/status/810201293151145984

See also: http://machinelearningmastery.com/popular-deep-learning-libraries

@jamieinfinity

Keras API is an intuitive high-level abstraction

Model setup

```
num_hidden_units = 25

model = Sequential()
model.add(Dense(num_hidden_units, input_dim=num_pixels, init='normal', activation='relu'))
model.add(Dense(num_classes, init='normal', activation='softmax'))
model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['accuracy'])
```

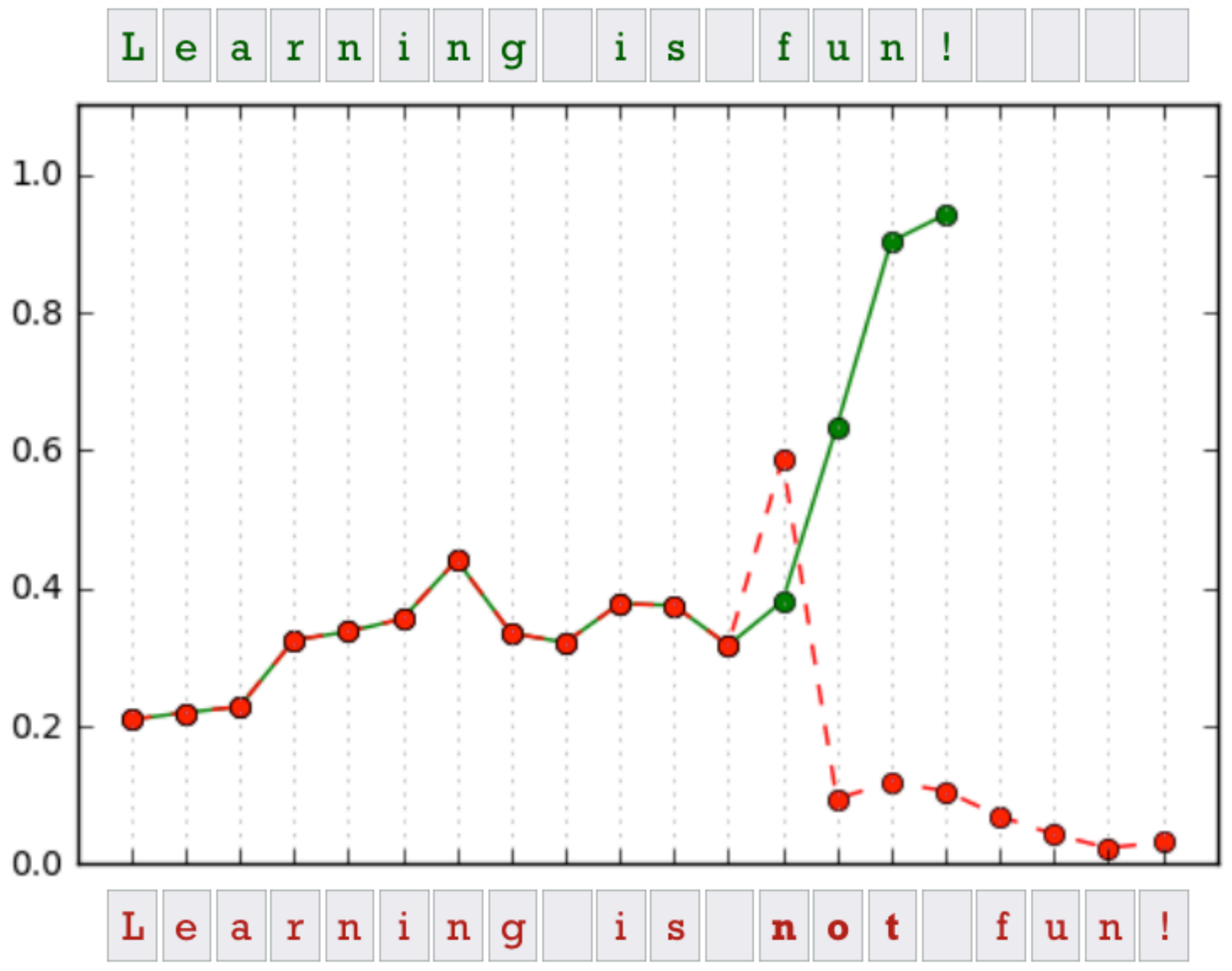
Training

```
model.fit(X_train, y_train, validation_data=(X_validation, y_validation),
          nb_epoch=15, batch_size=100)
```

Tutorial Applications of RNNs

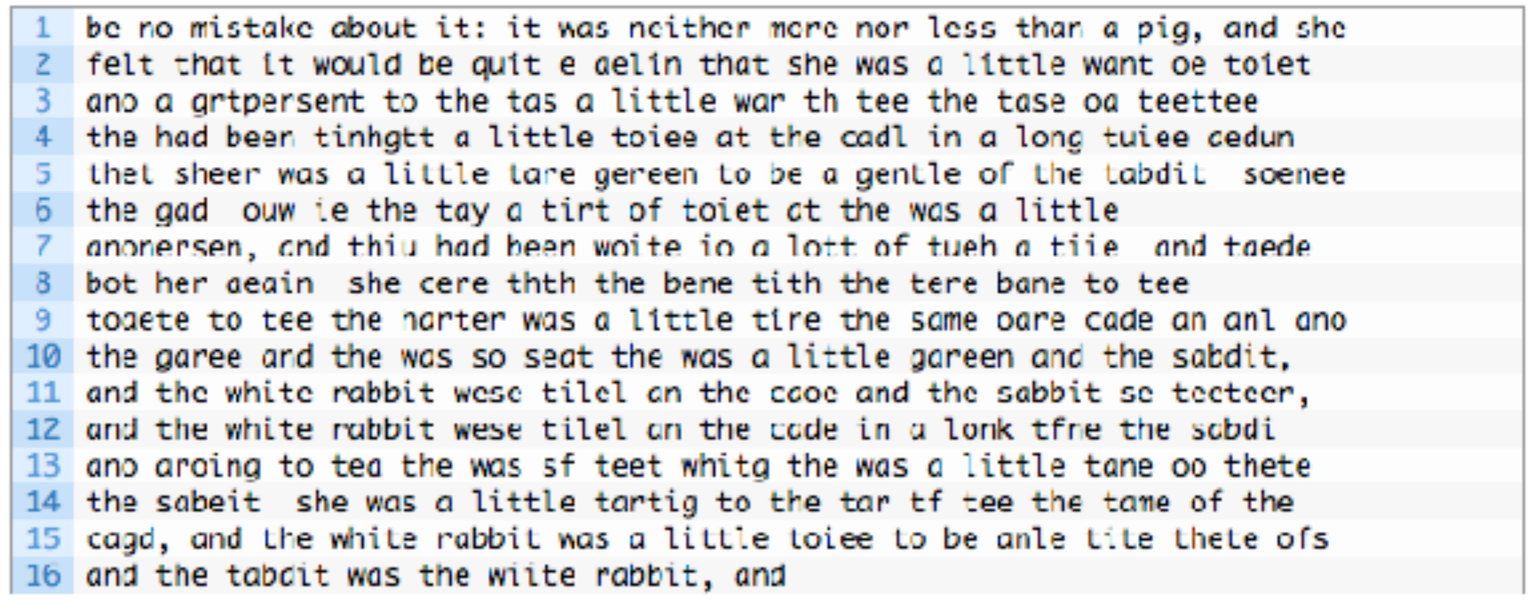
Machine Learning Mastery
machinelearningmastery.com

Sequence Classification with LSTMs



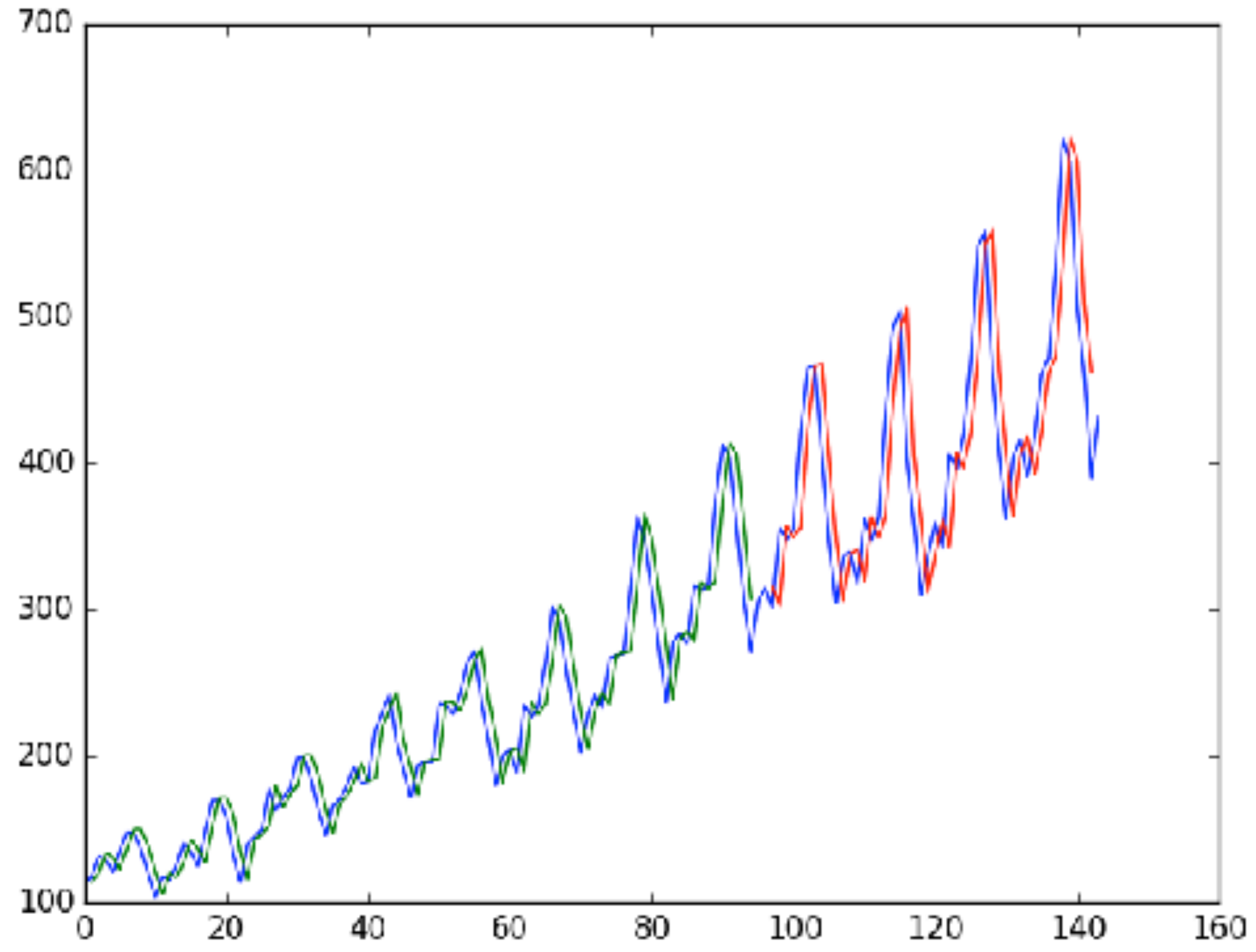
machinelearningmastery.com/sequence-classification-lstm-recurrent-neural-networks-python-keras/

Text Generation with LSTMs


be no mistake about it: it was neither more nor less than a pig, and she
felt that it would be quit e aelin that she was a little want oe toiet
ano a grtpersent to the tas a little war th tee the tase oa teettee
the had been tinhgtt a little toiee at the cadl in a long tuiee cedun
thel sheer was a little tare gereen to be a gentle of the tabdit soenee
the gad ouw ie the tay a tirt of toiet ct the was a little
anonersen, and thiu had been woite io a lott of tueh a tiie and taede
bot her aeain she cere thth the bene tith the tere bane to tee
toaete to tee the harter was a little tire the same oare cade an anl ano
the garee and the was so seat the was a little gareen and the sabdit,
and the white rabbit wese tilel an the cooe and the sabbit se teeteer,
and the white rubbit wese tilel an the cade in a lonk tfre the sabdi
ano aroing to tea the was sf teet whitg the was a little tane oo thete
the sabeit she was a little tartig to the tar tf tee the tame of the
cagd, and the white rabbit was a little toiee to be anle tite thete ofs
and the tabdit was the wiite rabbit, and

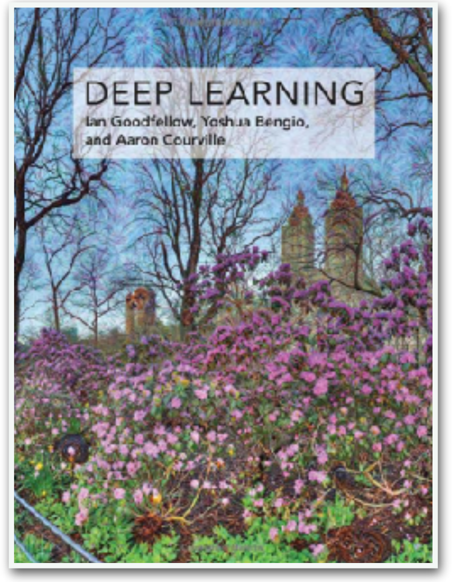
machinelearningmastery.com/ext-generation-lstm-recurrent-neural-networks-python-keras/

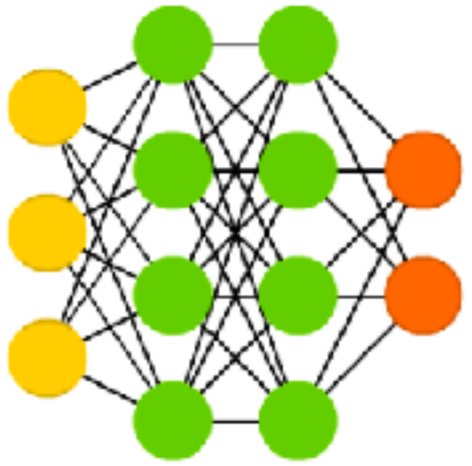
Time Series Prediction with LSTMs



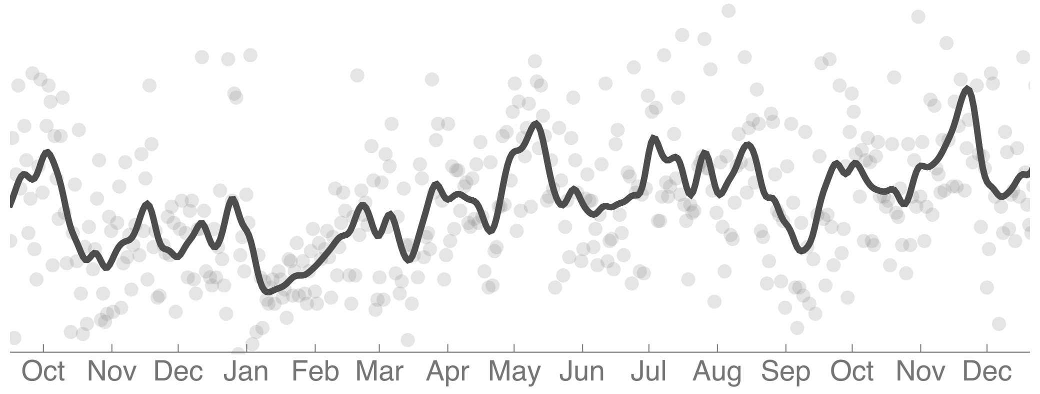
machinelearningmastery.com/time-series-prediction-with-deep-learning-in-python-with-keras/

Learning about deep learning

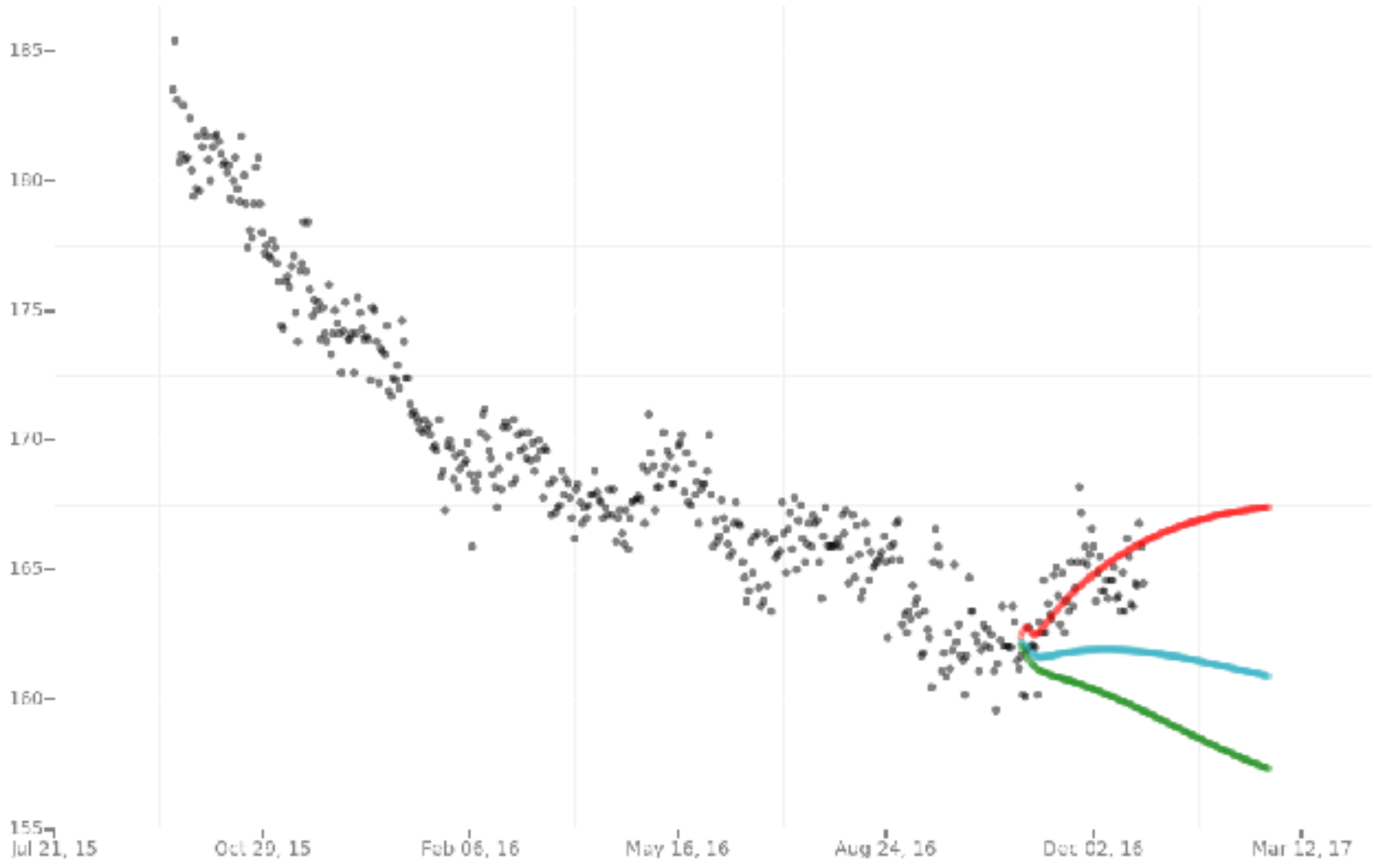




Collecting fitness data



Build an RNN predictive model of weight change



I’m a self tracker

physical activity

heart rate

computer & mobile usage

geolocation

weight / body composition

television & podcasts

auto trips

food & water consumption

finances

environment sensors

sleep

tweets

I’m a self tracker

physical activity

heart rate

computer & mobile usage

geolocation

weight / body composition

television & podcasts

auto trips

food & water consumption

finances

environment sensors

sleep

tweets

I'm a self tracker

physical activity

Fitbit

weight / body composition

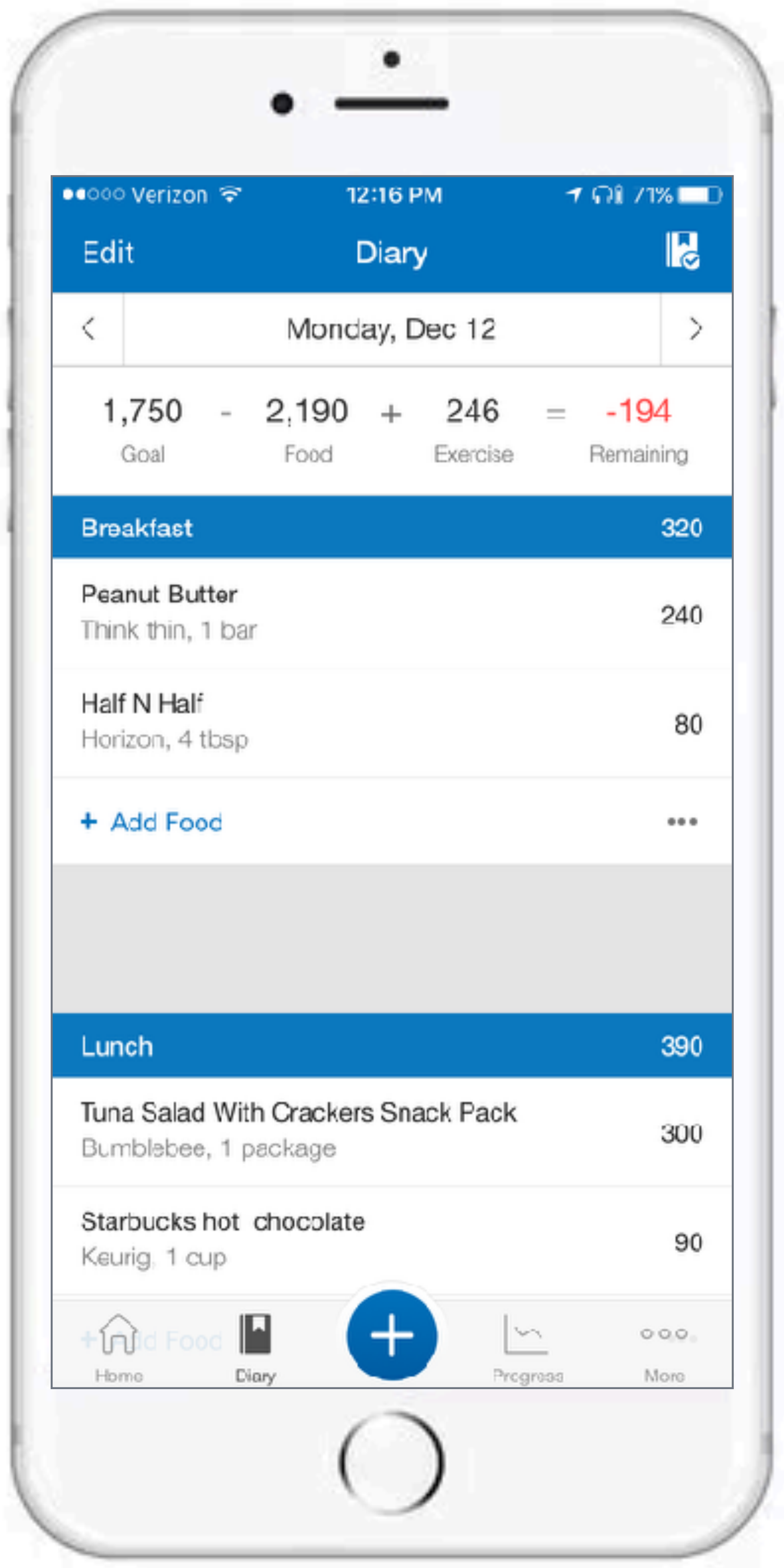
Withings

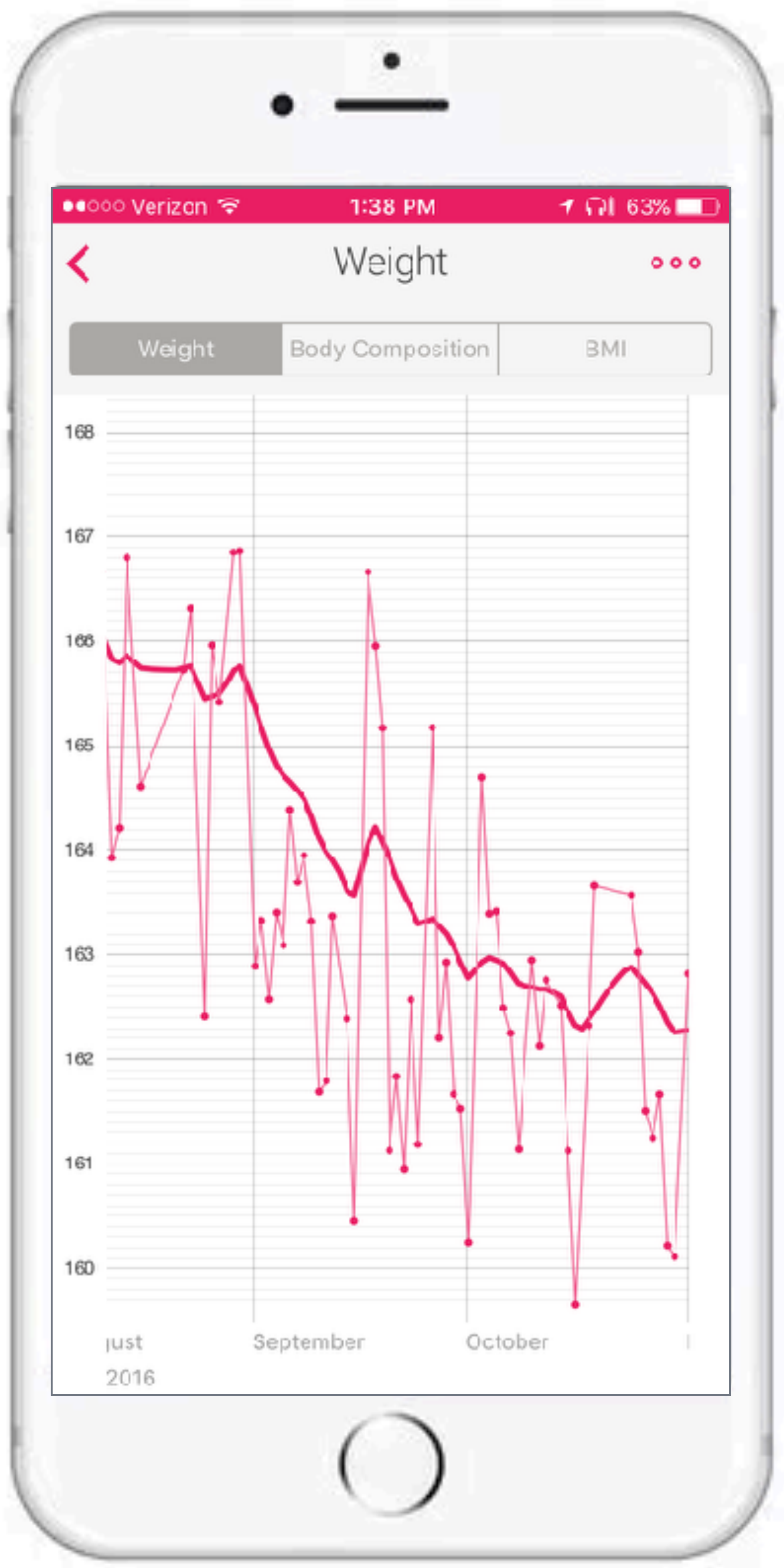
food & water consumption

MyFitnessPal

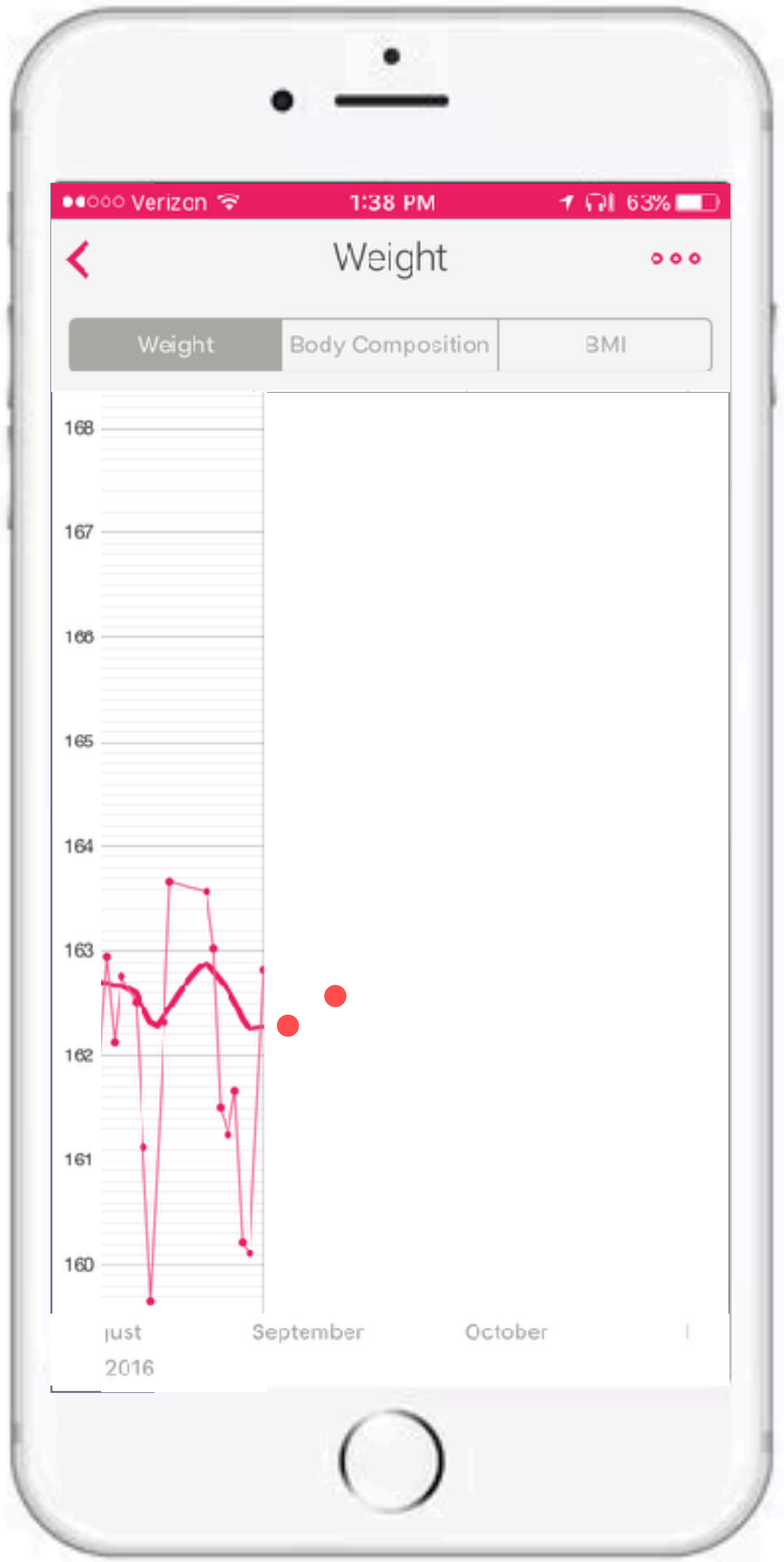

Weight (lbs)
weight
185
180
175
170
165
160
155
Steps
steps
30 000
25 000
20 000
15 000
10 000
5000
0
Calories
calories
3500
3000
2500
2000
1500
Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan Feb Mar Apr May Jun Jul Aug Sep Oct Nov Dec Jan
2013
2014
2015
2016
2017

Descriptive Analytics



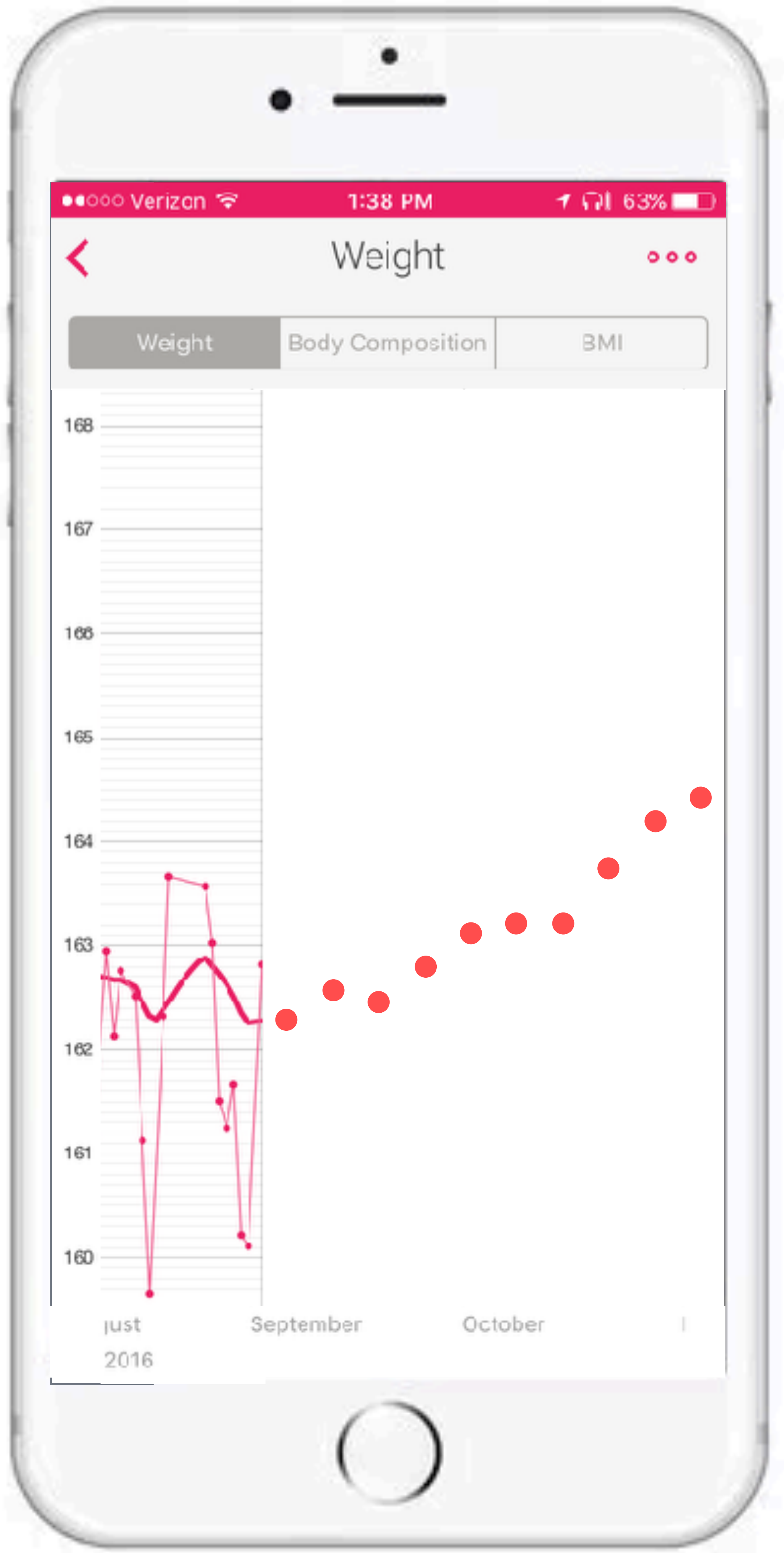


Predictive Analytics



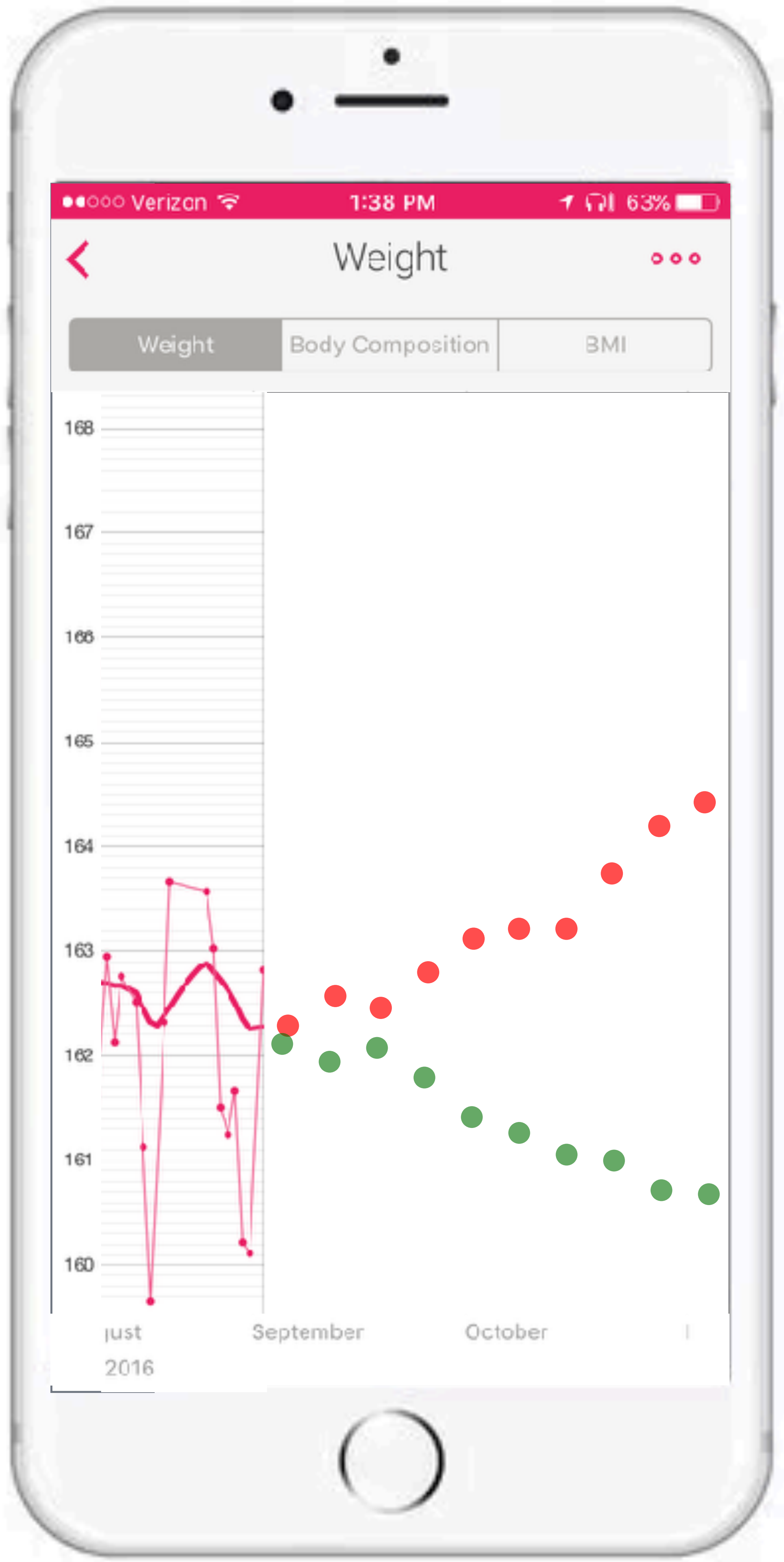
- forecast based on current behavior

Predictive Analytics



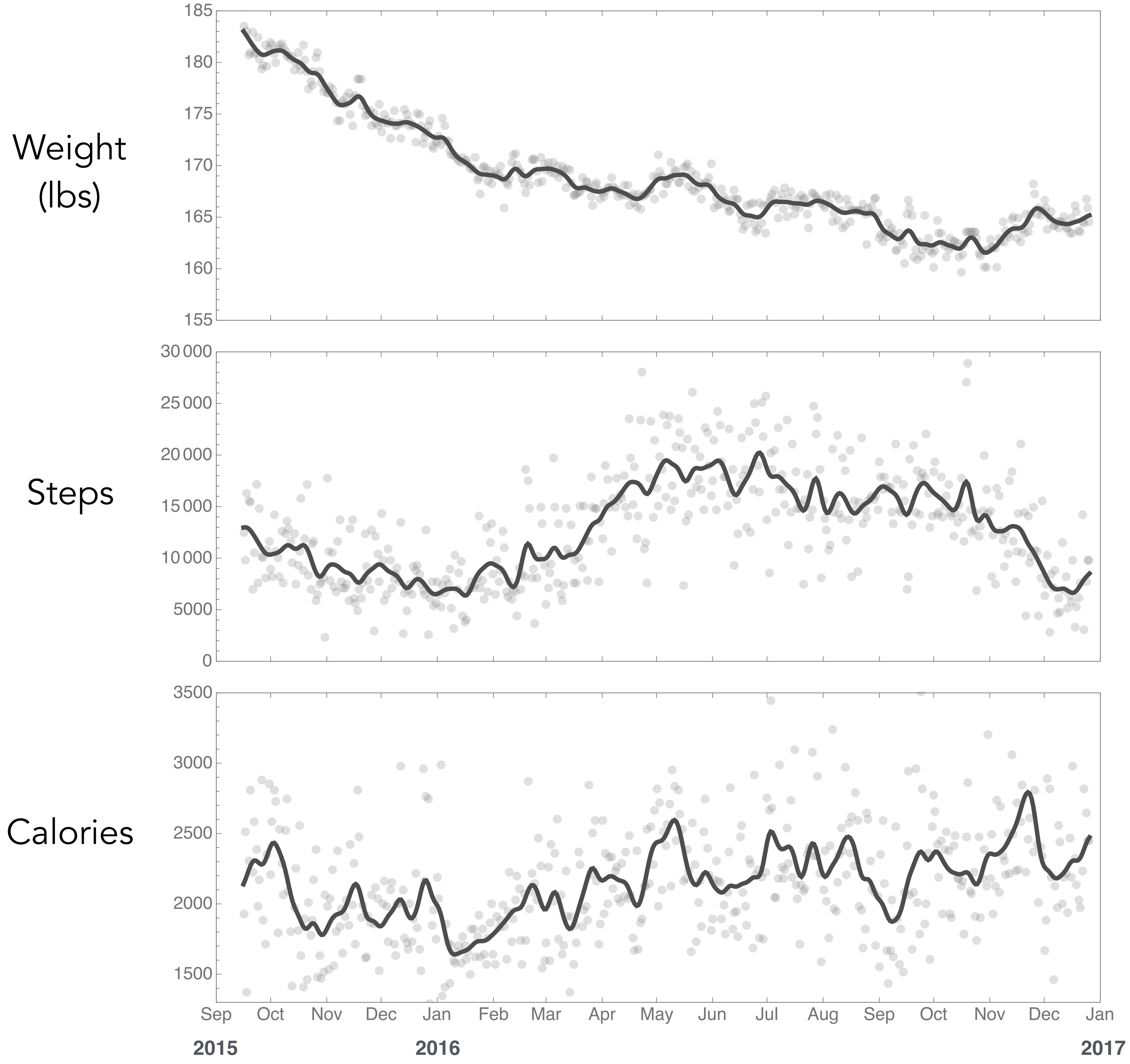
● forecast based on current behavior

Prescriptive Analytics



forecast based on current behavior

'What If' scenario
e.g. 5000 steps, eat 1900 calories



Challenge:

Given this historical data, come up with a way to predict daily **weight change** ΔW as a function of **calories** and **steps**.

$$\Delta W = W_{i+1} - W_i$$

$$\Delta W(C_i, S_i) = ?$$

Energy Balance Rate Equations

One appealing approach is to develop a **rate equation** model based on **energy balance** and physical processes in the body.

The Dynamics of Human Body Weight Change

Carson C. Chow*, Kevin D. Hall

Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland, United States of America

Abstract

An imbalance between energy intake and energy expenditure will lead to a change in body weight (mass) and body composition (fat and lean masses). A quantitative understanding of the processes involved, which currently remains lacking, will be useful in determining the etiology and treatment of obesity and other conditions resulting from prolonged energy imbalance. Here, we show that a mathematical model of the macronutrient flux balances can capture the long-term dynamics of human weight change; all previous models are special cases of this model. We show that the generic dynamic behavior of body composition for a clamped diet can be divided into two classes. In the first class, the body composition and mass are determined uniquely. In the second class, the body composition can exist at an infinite number of possible states. Surprisingly, perturbations of dietary energy intake or energy expenditure can give identical responses in both model classes, and existing data are insufficient to distinguish between these two possibilities. Nevertheless, this distinction has important implications for the efficacy of clinical interventions that alter body composition and mass.

Citation: Chow CC, Hall KD (2008) The Dynamics of Human Body Weight Change. PLoS Comput Biol 4(3): e1000045. doi:10.1371/journal.pcbi.1000045

Editor: Philip E. Bourne, University of California San Diego, United States of America

Received November 26, 2007; **Accepted** February 28, 2008; **Published** March 28, 2008

Funding: This research was supported by the intramural research program of the NIH/NIDDK.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: carsonc@mail.nih.gov

$$\rho_F \frac{dF}{dt} = I_F - f_F E$$

fat mass

$$\rho_G \frac{dG}{dt} = I_C - f_C E$$

glycogen mass

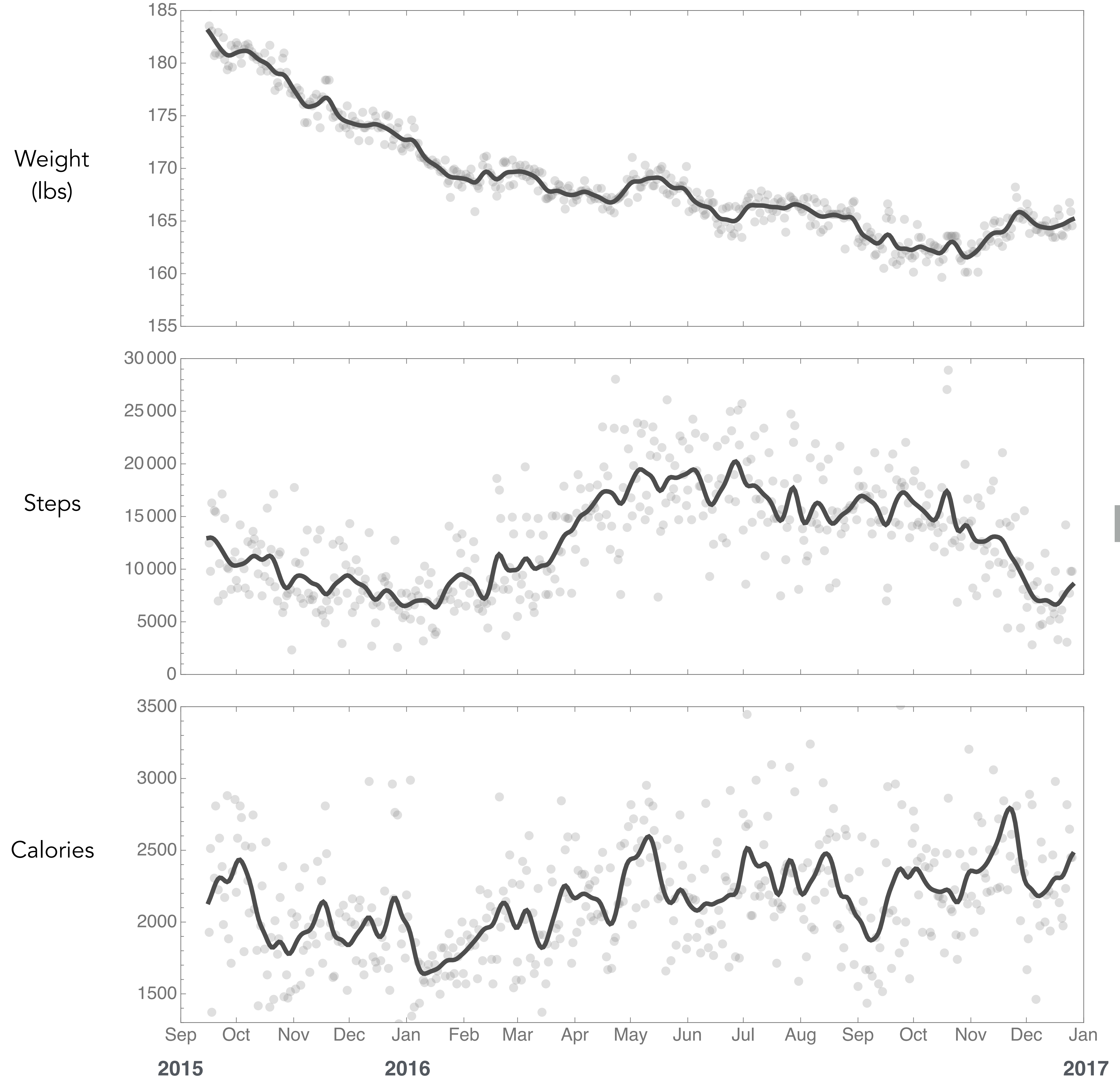
$$\rho_P \frac{dP}{dt} = I_P - (1 - f_F - f_C) E$$

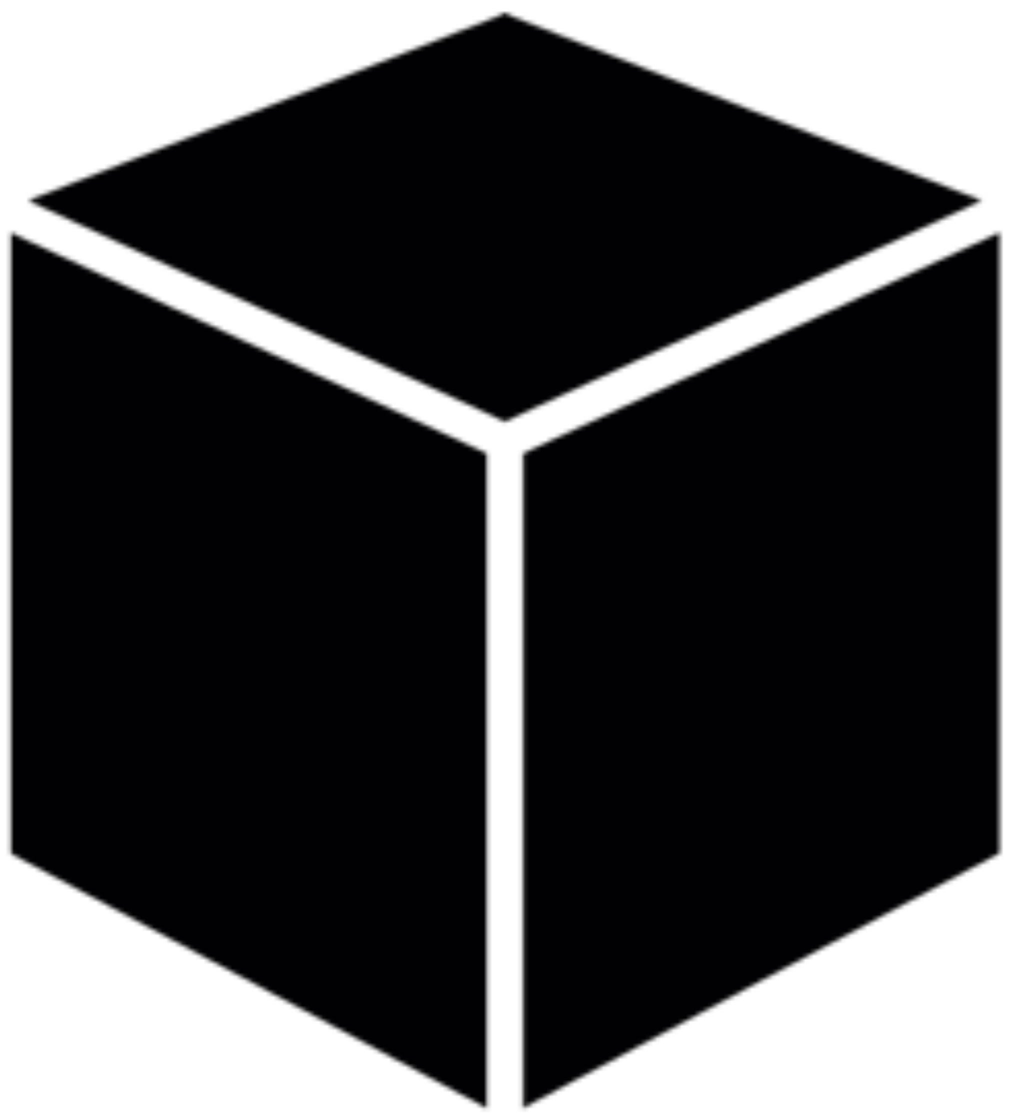
protein mass

References:

- Chow CC, Hall KD (2008) *The Dynamics of Human Body Weight Change*. PLoS Comput Biol 4(3): e1000045.
- Diana M. Thomas , Corby K. Martin , Steven Heymsfield , Leanne M. Redman , Dale A. Schoeller & James A. Levine (2011) *A simple model predicting individual weight change in humans*, Journal of Biological Dynamics, 5:6, 579-599.

A Machine Learning Approach





$$\Delta W(C_i, S_i)$$

Setting up the Problem

	Weight	W_1	W_2	W_3	W_4	W_5	W_6	W_7	W_8	W_9	W_{10}	W_{11}	W_{12}
Features	**Steps**	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}
	Calories	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}
	Day	1	2	3	4	5	6	7	8	9	10	11	12

Setting up the Problem

Features	$\vec{x}_1$	$\vec{x}_2$	$\vec{x}_3$	$\vec{x}_4$	$\vec{x}_5$	$\vec{x}_6$	$\vec{x}_7$	$\vec{x}_8$	$\vec{x}_9$	$\vec{x}_{10}$	$\vec{x}_{11}$	$\vec{x}_{12}$	∘∘∘
Day	1	2	3	4	5	6	7	8	9	10	11	12	

Setting up the Problem

Problem Statement: Given data for n days, I want to predict what my weight will be for the day $n + 1$. For example, taking $n = 5$:

$$\vec{x}_1 \quad \vec{x}_2 \quad \vec{x}_3 \quad \vec{x}_4 \quad \vec{x}_5 \quad \xrightarrow{\text{predict}} \quad W_6$$

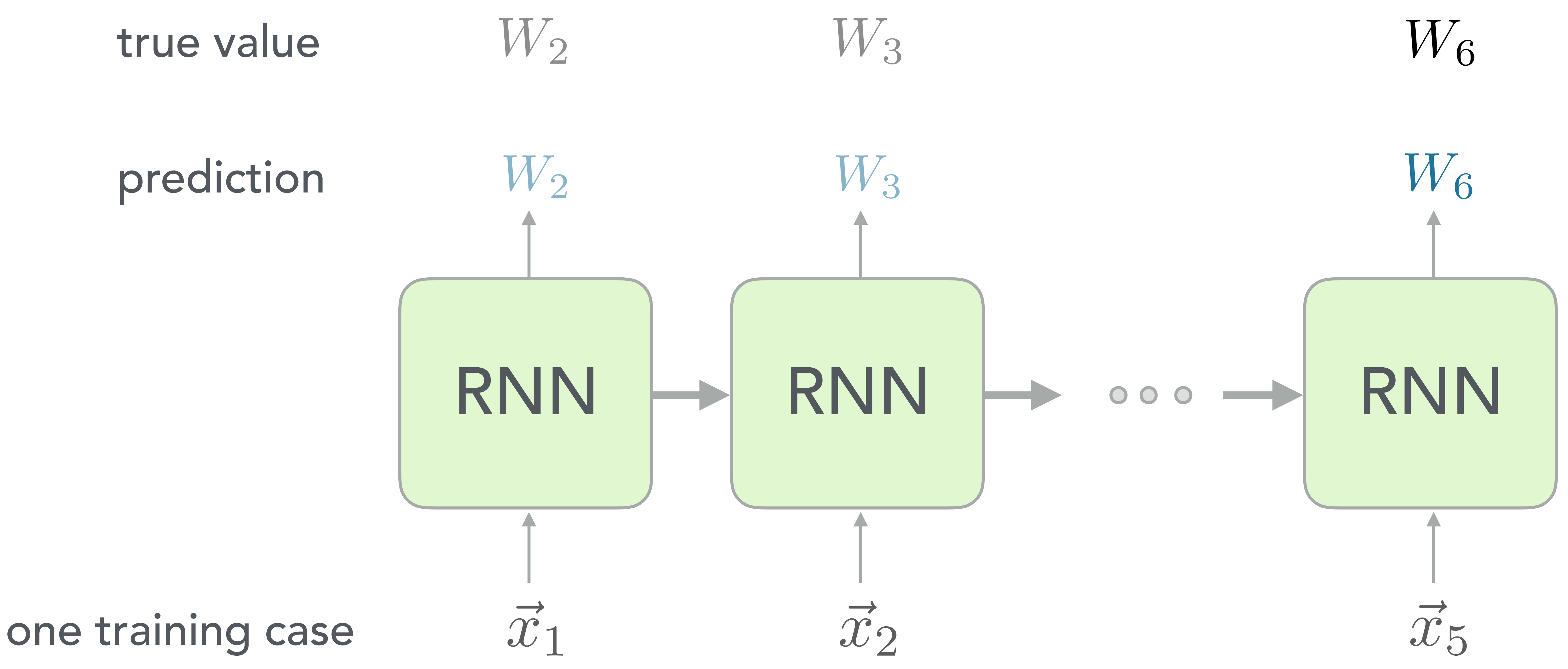
Splice data into sequences

We thus need to splice the data into sequences of length n, e.g.

sequence of feature vectors						target
$\vec{x}_1$	$\vec{x}_2$	$\vec{x}_3$	$\vec{x}_4$	$\vec{x}_5$	$\xrightarrow{\text{predict}}$	W_6
$\vec{x}_2$	$\vec{x}_3$	$\vec{x}_4$	$\vec{x}_5$	$\vec{x}_6$	$\xrightarrow{\text{predict}}$	W_7
$\vec{x}_3$	$\vec{x}_4$	$\vec{x}_5$	$\vec{x}_6$	$\vec{x}_7$	$\xrightarrow{\text{predict}}$	W_8
$\vec{x}_4$	$\vec{x}_5$	$\vec{x}_6$	$\vec{x}_7$	$\vec{x}_8$	$\xrightarrow{\text{predict}}$	W_9
$\vec{x}_5$	$\vec{x}_6$	$\vec{x}_7$	$\vec{x}_8$	$\vec{x}_9$	$\xrightarrow{\text{predict}}$	W_{10}

$\vdots$

Loss Function

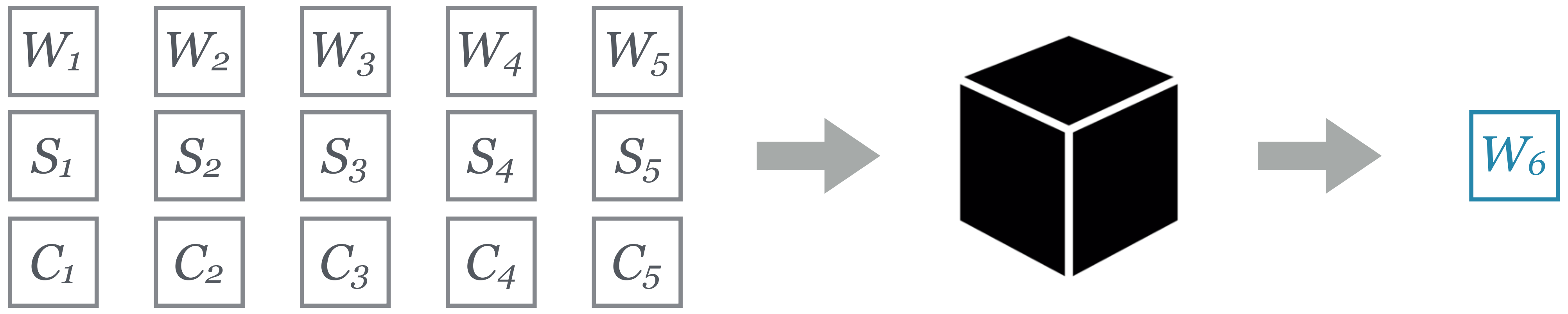


mean absolute error (MAE)

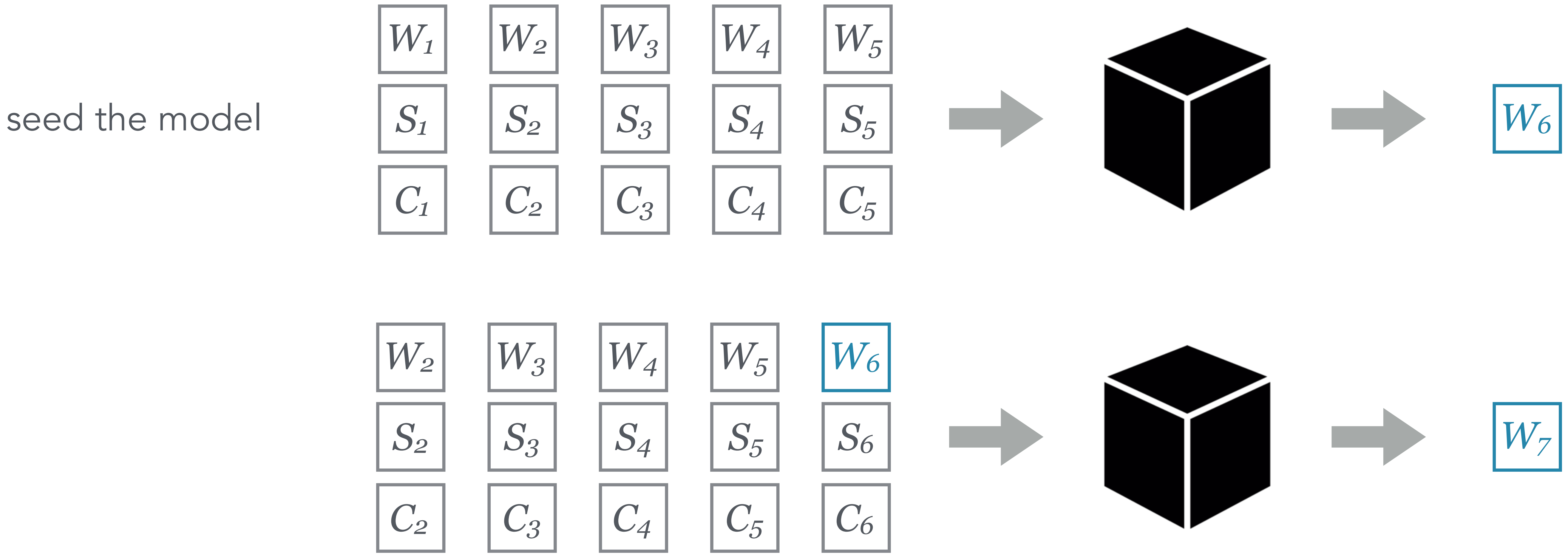
$$|W_6 - W_6|$$

Generating a Forecast

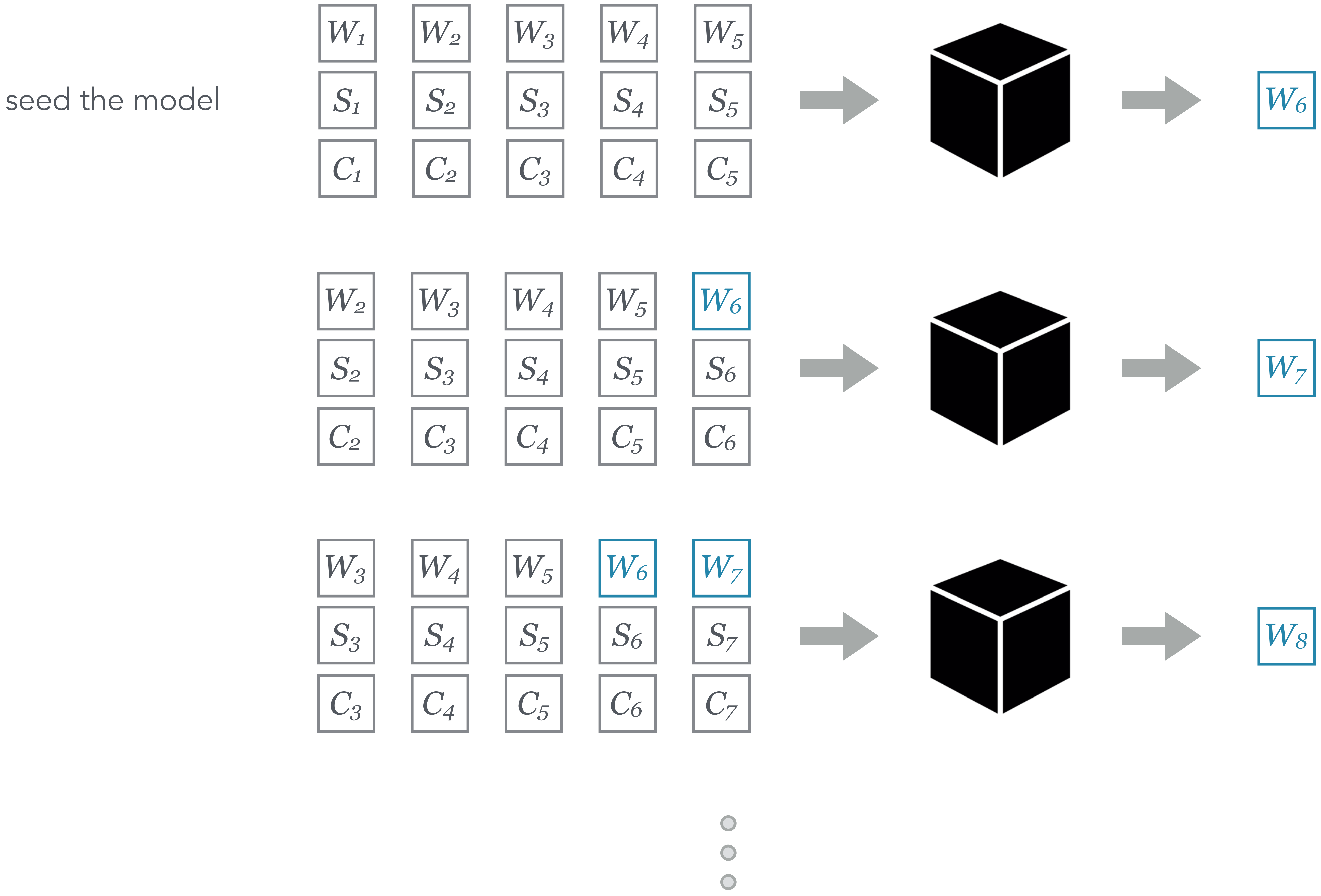
seed the model



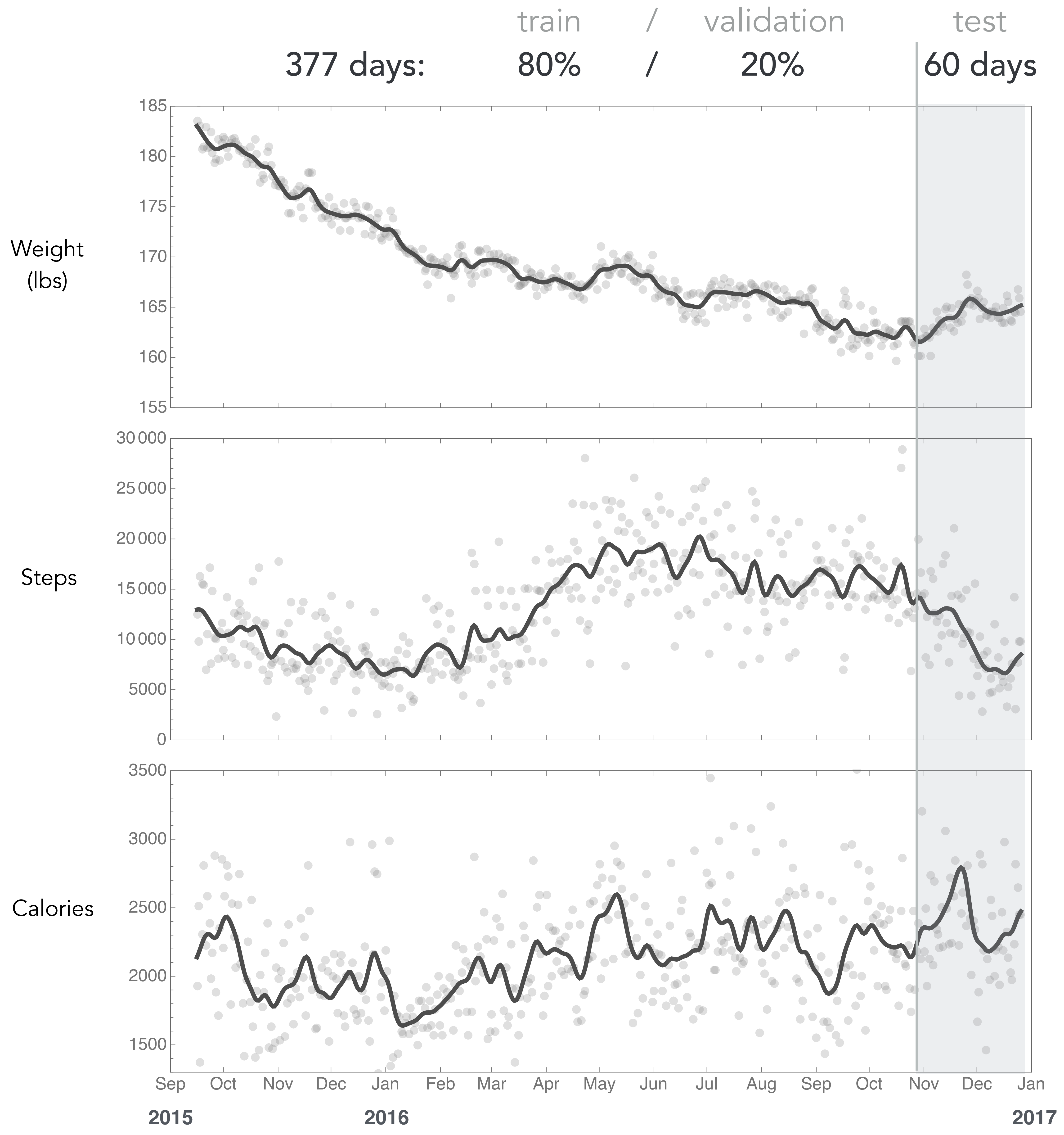
Generating a Forecast



Generating a Forecast



Training / Validation / Test



The Model Architecture in Keras

```
sequence_length = 30
loss_function = 'mean_absolute_error'
optimization_method = 'adam'
lstm_size = 30
lstm2_size = 30
dropout_rate = 0.3
dropout_rate2 = 0.5
batch_size = 10
number_of_epochs = 150
early_stopping_patience = 5
```

optimal **hyperparameters** obtained from grid search

```
layer_input = Input(shape=(sequence_length, 3))
layer_lstm = LSTM(lstm_size, return_sequences=True)(layer_input)
layer_dropout = Dropout(dropout_rate)(layer_lstm)
layer_lstm2 = LSTM(lstm2_size, return_sequences=True)(layer_dropout)
layer_dropout2 = Dropout(dropout_rate2)(layer_lstm2)
layer_output = TimeDistributed(Dense(1))(layer_dropout2)
```

input
LSTM
dropout
LSTM
dropout
dense

fitting parameters: 11,431

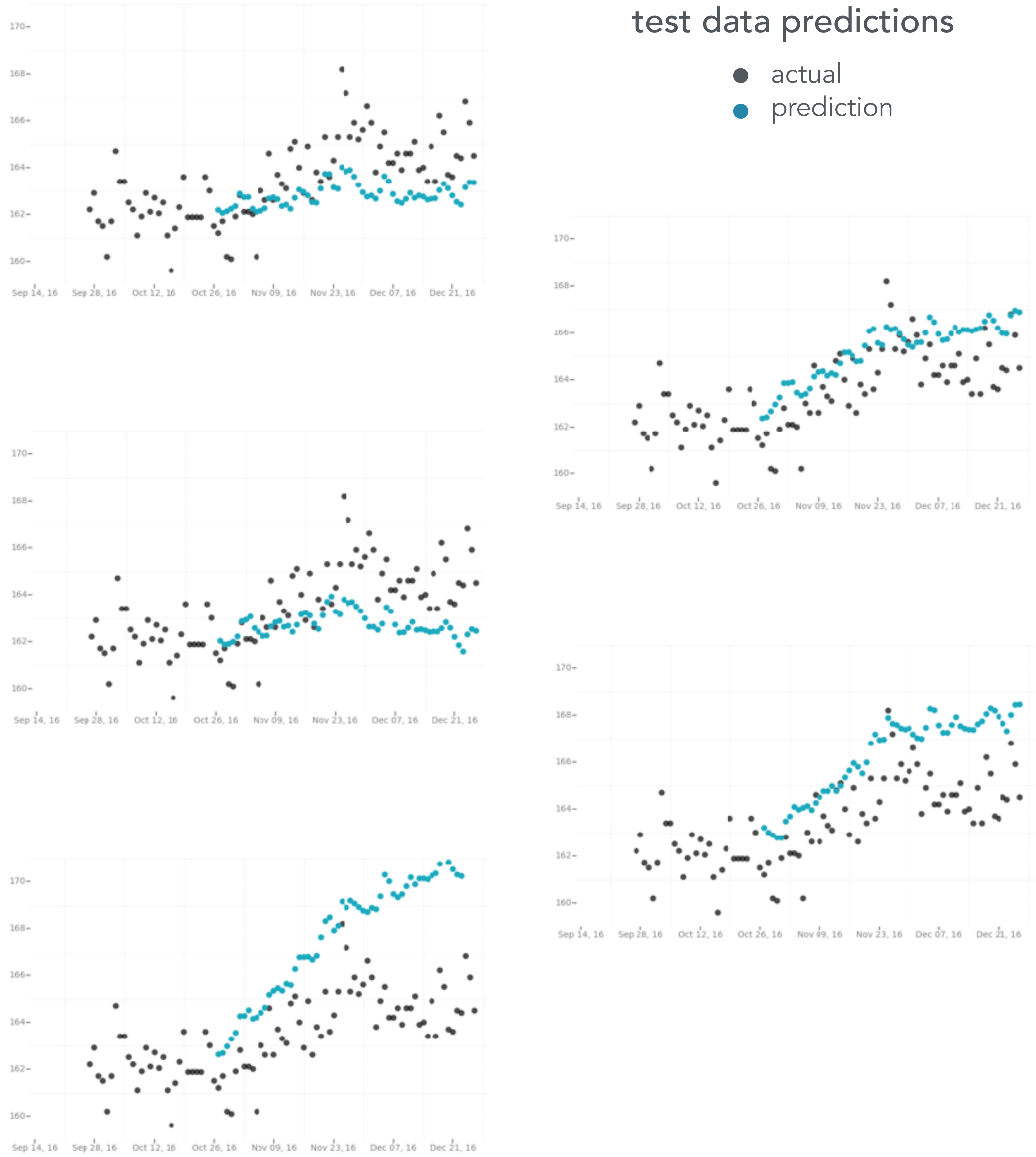
```
model = Model(input=[layer_input], output=[layer_output])
model.compile(loss=loss_function, optimizer=optimization_method)

early_stopping = EarlyStopping(monitor='val_loss', patience=early_stopping_patience)

model.fit(X_train, Y_train, validation_data=(X_val, Y_val),
          batch_size=batch_size,
          nb_epoch=number_of_epochs, verbose=verboseness, callbacks=[early_stopping])
```

k-fold Cross Validation Results

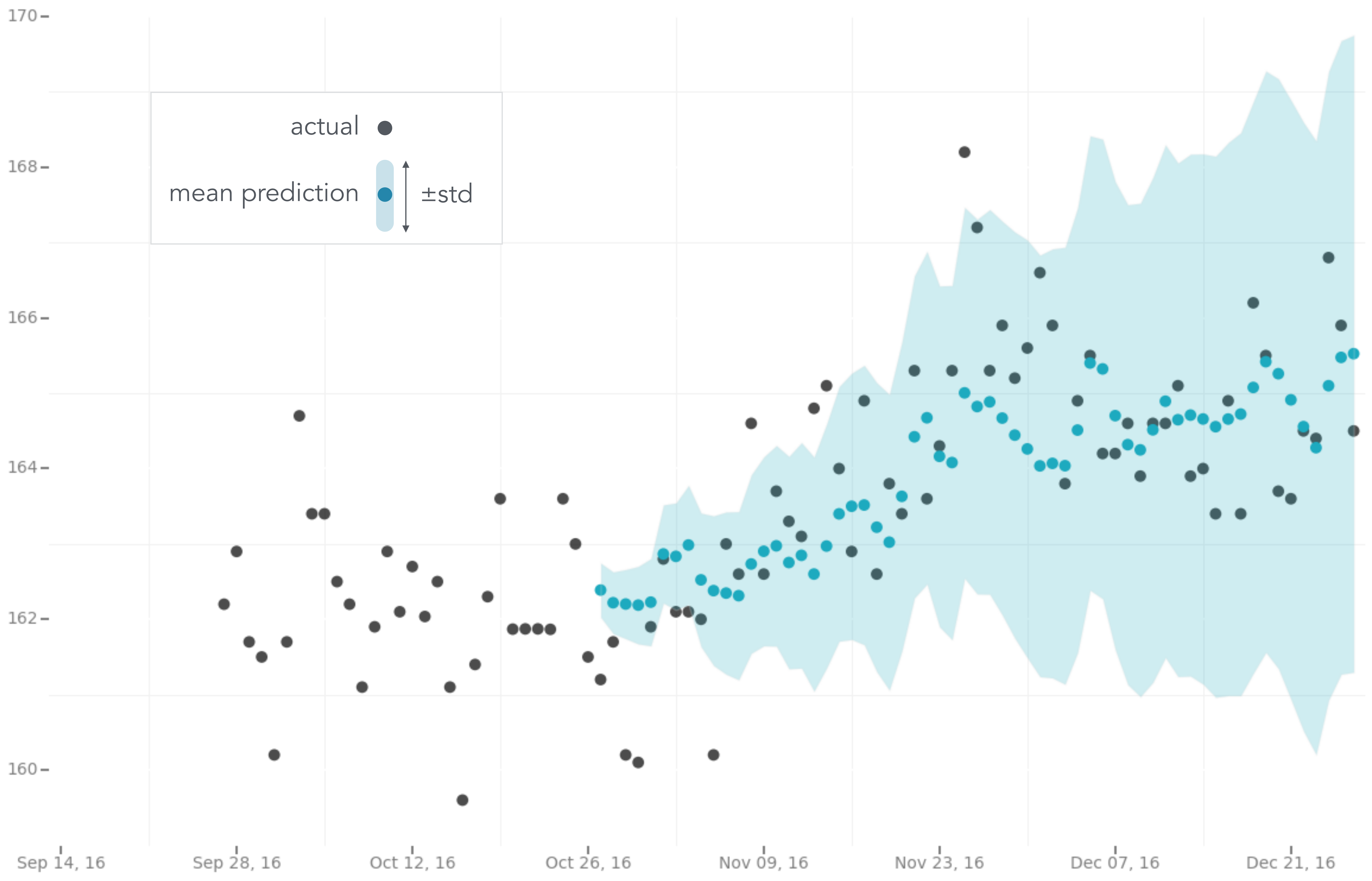
fold	mean absolute error MAE_{30}
1	0.766 lbs
2	0.755 lbs
3	0.778 lbs
4	0.750 lbs
5	0.777 lbs

k-fold Cross Validation Results

fold	mean absolute error MAE_{30}	MAE on test set
1	0.766 lbs	1.53 lbs
2	0.755 lbs	1.38 lbs
3	0.778 lbs	1.68 lbs
4	0.750 lbs	2.25 lbs
5	0.777 lbs	3.74 lbs



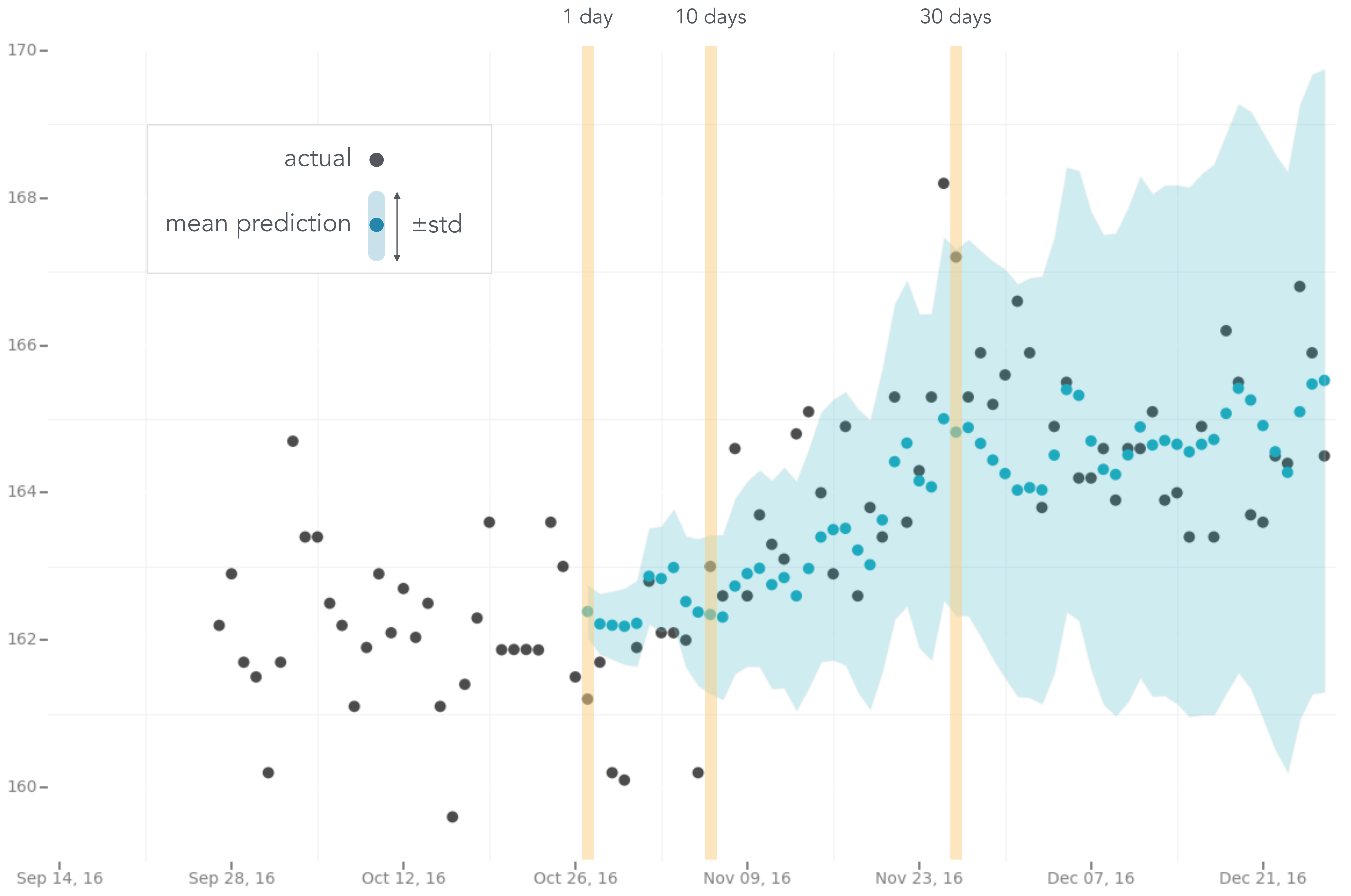
Wisdom of the Crowds



Train 100 models on independent random 80/20 splits of the training/val dataset.

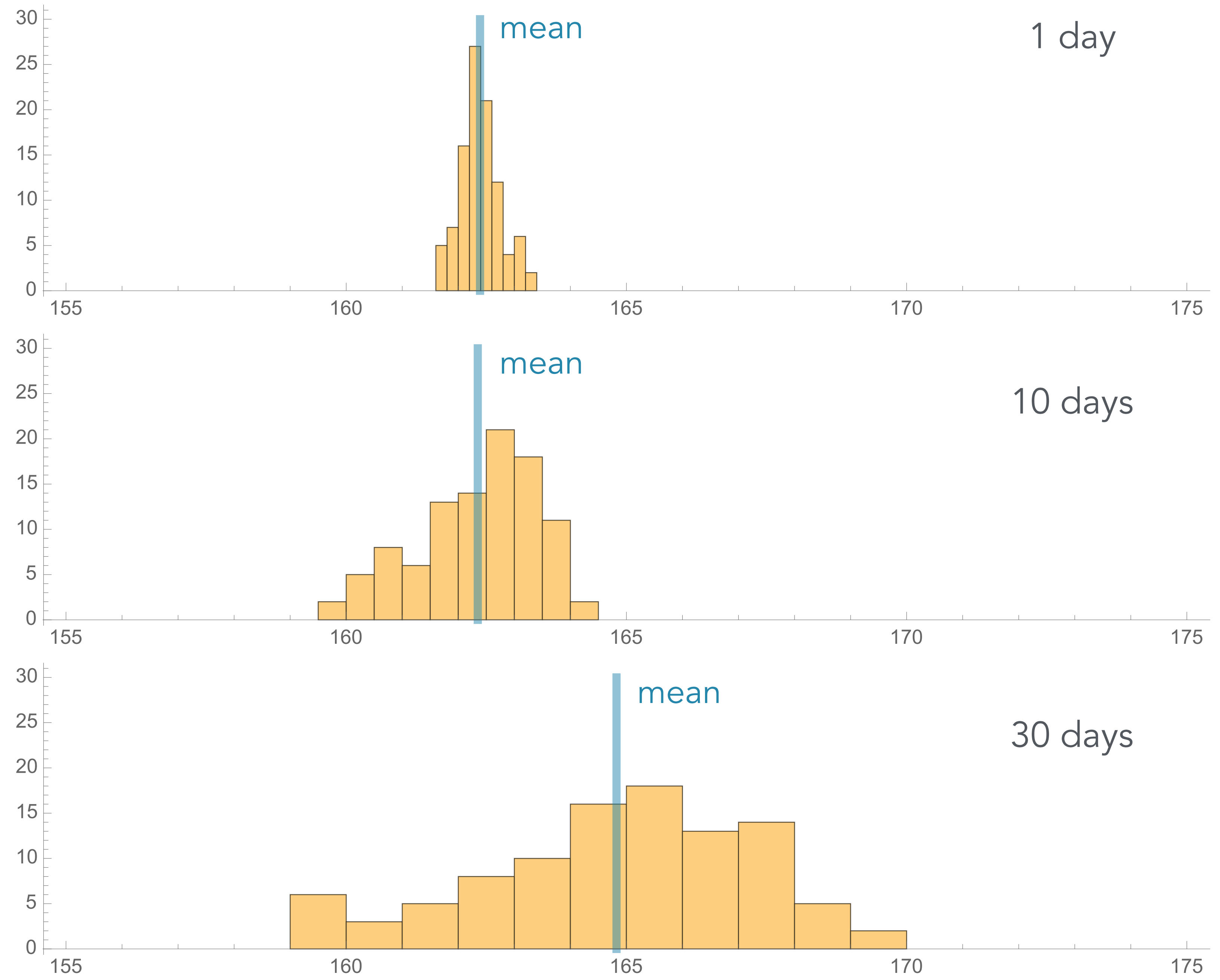
Take average of predictions on the test data.

Wisdom of the Crowds



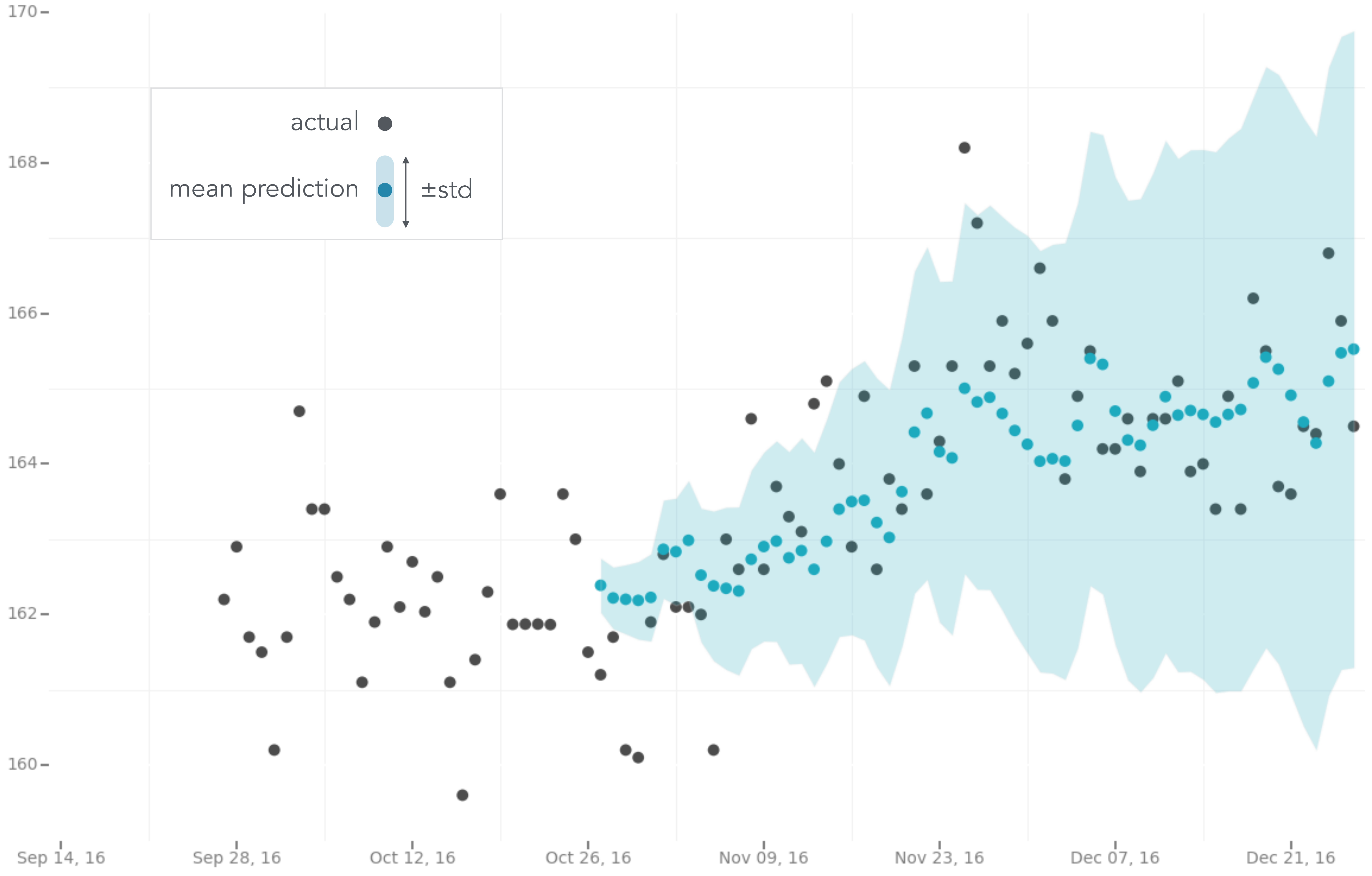
Train 100 models on independent random 80/20 splits of the training/val dataset.

Take average of predictions on the test data.



@jamieinfinity

Wisdom of the Crowds



Train 100 models on independent random 80/20 splits of the training/val dataset.

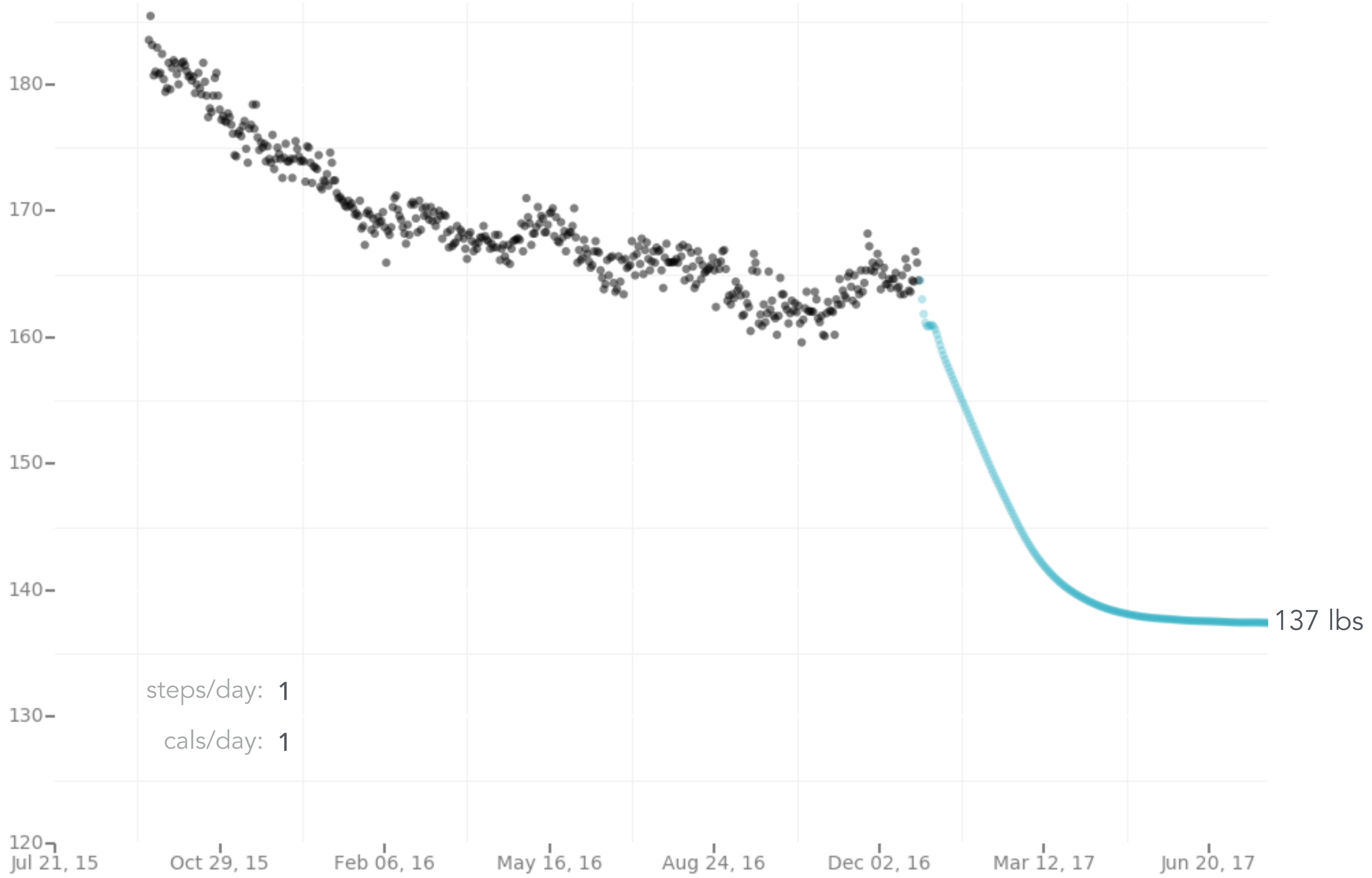
Take average of predictions on the test data.

MAE on test data

0.92 lbs

Model has physically intuitive behavior

Realistic asymptotic limit of starvation: If I decrease calories to 0, the model predicts a leveling off to 137 lbs.



According to Withings scale:

lean mass: 127 lbs

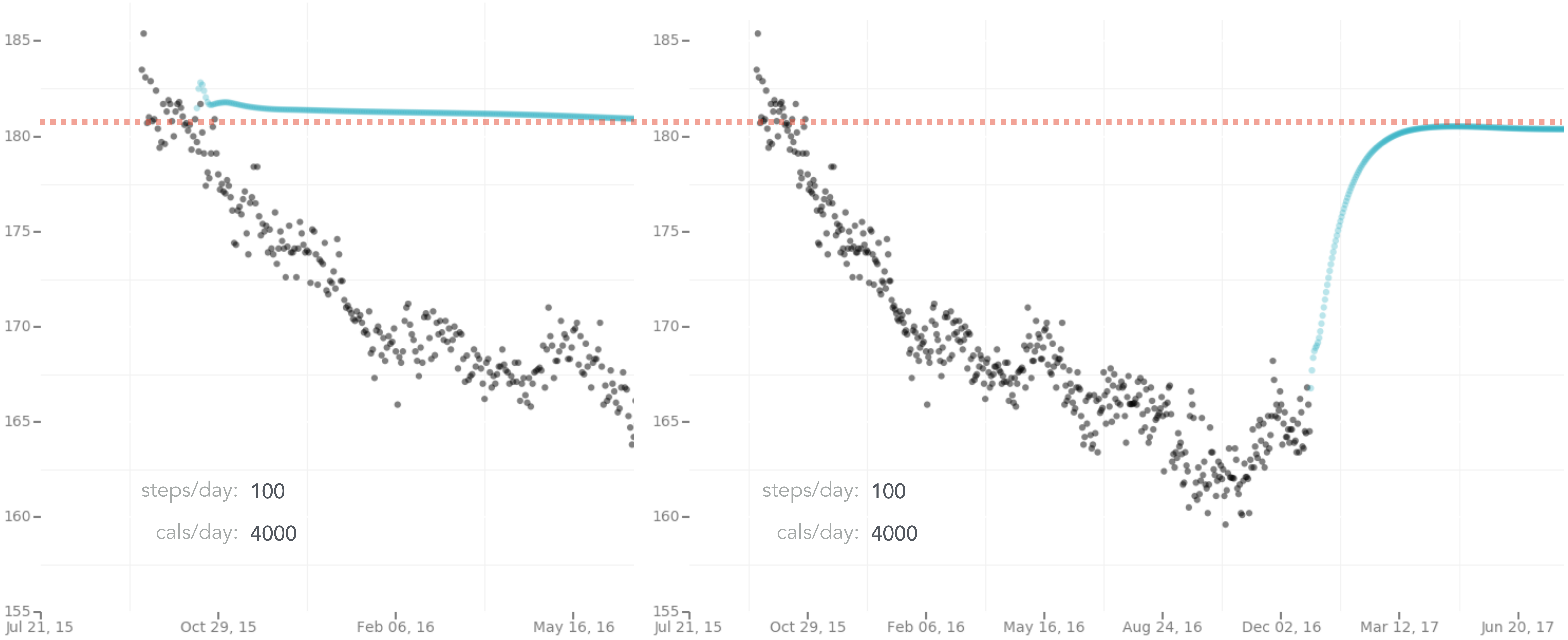
fat mass: 32 lbs

bone mass: 6 lbs

Model has physically intuitive behavior

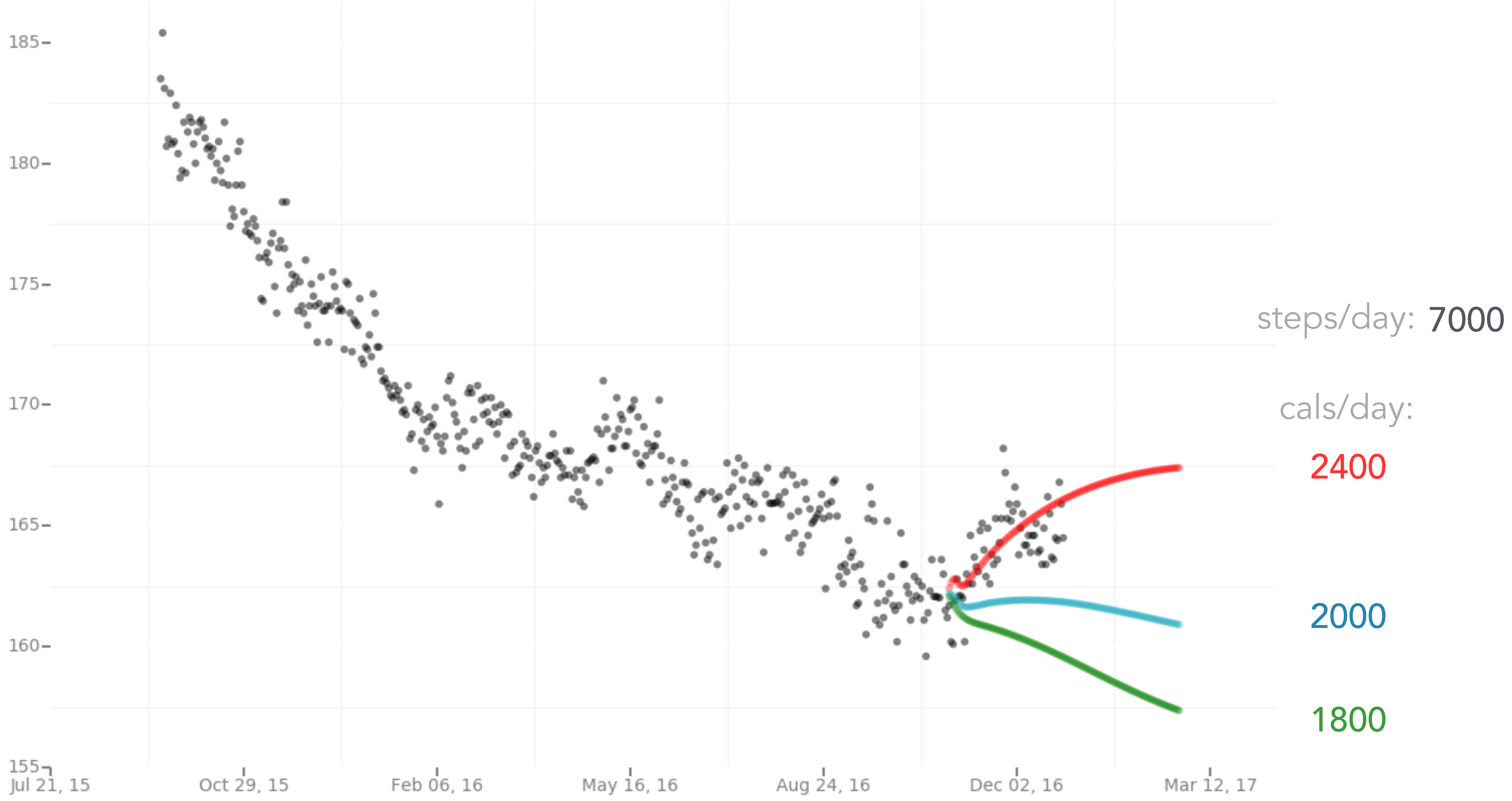
Internal self-consistency:

- To maintain a weight of 180 lbs, I would have needed to eat a 4000 cal/day diet.
- If today I were to eat a 4000 cal/day diet, my weight would increase up to 180 lbs.

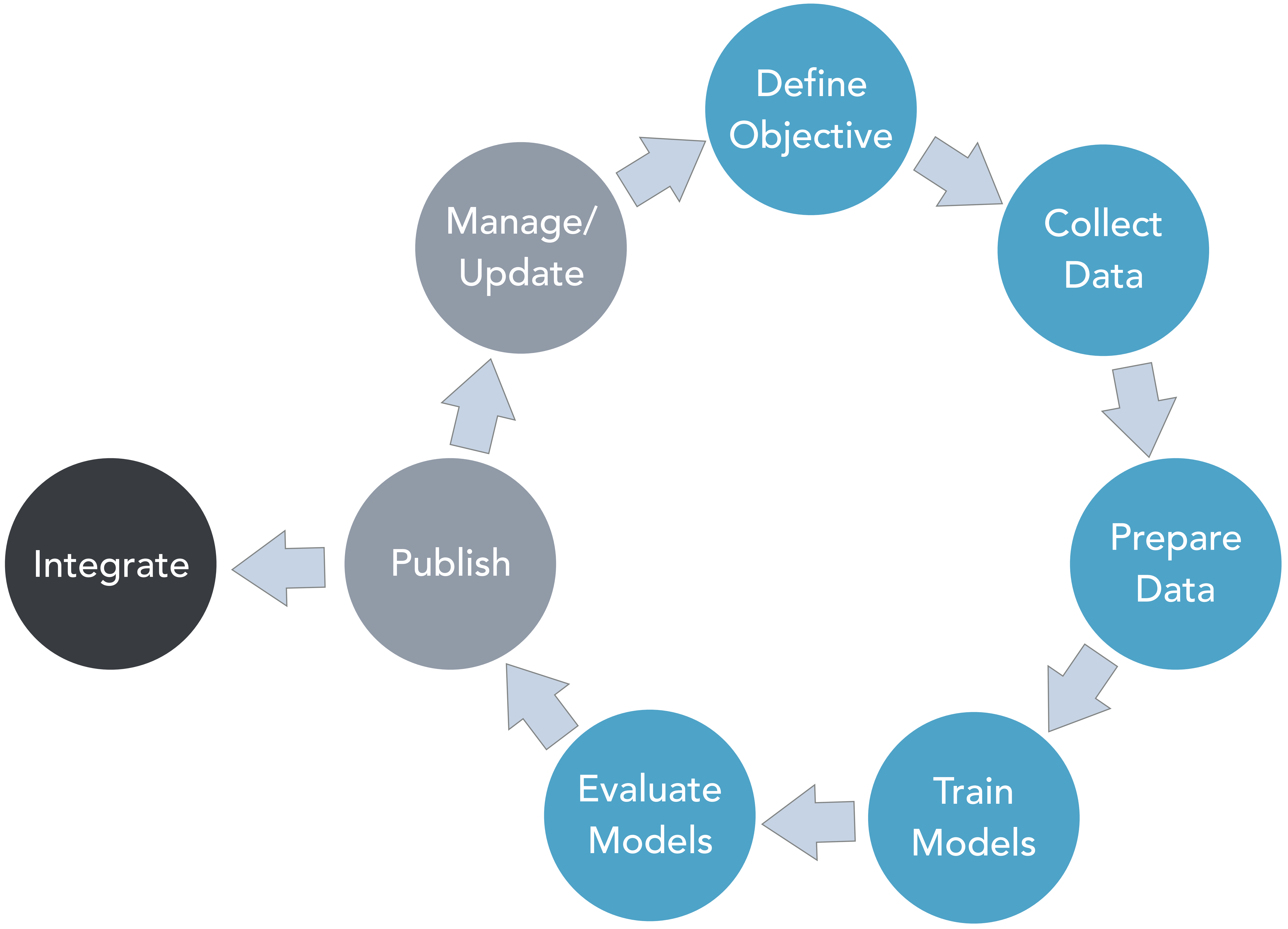


What-if Scenarios

As winter approached, I knew my step count would decrease. What could I expect under different scenarios?



Next step: Give it to a user (me!)



Future extensions of model

- **benchmark** against energy balance rate equation model
- include **more features**, e.g. activity & calorie breakdowns
- **data augmentation** with noisy data (based on uncertainties in measurements)